CORRECTED VERSION*

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5 : C12P 19/34

A1

(11) International Publication Number: WO 98/03673

(43) International Publication Date: 29 January 1998 (29.01.98)

(21) International Application Number: PCT/US97/12195

(22) International Filing Date: 15 July 1997 (15.07.97)

(30) Priority Data:
60/022,535 19 July 1996 (19.07.96) US

(71) Applicants: CORNELL RESEARCH FOUNDATION, INC. [US/US]; Suite 105, 20 Thornwood Drive, Ithaca, NY 14850 (US). PURDUE RESEARCH FOUNDATION [US/US]; Office of Technology Transfer, 1650 Engineering Administration Building, Room 328, West Lafayette, IN 47907-1650 (US).

(72) Inventors: BARANY, Francis; Apartment 12C, 450 E. 63rd Street, New York, NY 10021 (US). LUO, Jianying; Apartment #11, 120 Peterborough Street, Boston, MA 02215 (US). KHANNA, Marilyn; Apartment #4G, 500 East 63rd Street, New York, NY 10021 (US). BERGSTROM, Donald, E.; 3416 Hamilton Street, West Lafayette, IN 47906 (US).

(74) Agents: GOLDMAN, Michael, L. et al.; Nixon, Hargrave, Devans & Doyle LLP, Clinton Square, P.O. Box 1051, Rochester, NY 14603 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ARIPO patent (GH, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: HIGH FIDELITY DETECTION OF NUCLEIC ACID DIFFERENCES BY LIGASE DETECTION REACTION

(57) Abstract

Ligase detection reaction is utilized to distinguish minority template in the presence of an excess of normal template with a thermostable ligase. This process can be carried out with a mutant ligase, thermostable ligase, or a modified oligonucleotide probe. This procedure is particularly useful for the detection of cancer-associated mutations. It has the advantage of providing a quantitative measure of the amount or ratio of minority template.

*(Referred to in PCT Gazette No. 20/1998, Section II)

BNSDOCID: <WO___9803673A1_IA>

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

BNSDOCID: <WO___9803673A1_IA>

HIGH FIDELITY DETECTION OF NUCLEIC ACID DIFFERENCES BY LIGASE DETECTION REACTION

This application claims the benefit of U.S. Provisional Patent Application Serial No. 60/022,535, filed July 19, 1996.

This invention was developed with government funding under National Institutes of Health Grant No. GM41337-06. The U.S. Government may retain certain rights.

FIELD OF THE INVENTION

The present invention relates to the high fidelity detection of nucleic acid sequence differences using ligase detection reaction ("LDR"). One aspect of the present invention involves use of a ligase detection reaction to distinguish minority template in the presence of an excess of normal template with a thermostable ligase. Another aspect of the present invention relates to the use of a mutant ligase to carry out a ligase detection reaction. A third aspect of the present invention involves use of a modified oligonucleotide probe to carry out a ligase detection reaction.

BACKGROUND OF THE INVENTION

Multiplex Detection

Large-scale multiplex analysis of highly polymorphic loci is needed for practical identification of individuals, e.g., for paternity testing and in forensic science (Reynolds et al., Anal. Chem., 63:2-15 (1991)), for organ-transplant donor-recipient matching (Buyse et al., Tissue Antigens, 41:1-14 (1993) and Gyllensten et al., PCR Meth. Appl, 1:91-98 (1991)), for genetic disease diagnosis, prognosis, and pre-natal counseling (Chamberlain et al., Nucleic Acids Res., 16:11141-11156 (1988) and L. C. Tsui, Human Mutat., 1:197-203 (1992)), and the study of oncogenic mutations (Hollstein et al., Science, 253:49-53 (1991)). In addition, the cost-effectiveness of infectious disease diagnosis by nucleic acid analysis varies directly with the multiplex scale in panel testing. Many of these applications depend on the discrimination of single-base differences at a multiplicity of sometimes closely spaced loci.

A variety of DNA hybridization techniques are available for detecting the presence of one or more selected polynucleotide sequences in a sample containing

a large number of sequence regions. In a simple method, which relies on fragment capture and labeling, a fragment containing a selected sequence is captured by hybridization to an immobilized probe. The captured fragment can be labeled by hybridization to a second probe which contains a detectable reporter moiety.

Another widely used method is Southern blotting. In this method, a mixture of DNA fragments in a sample is fractionated by gel electrophoresis, then fixed on a nitrocellulose filter. By reacting the filter with one or more labeled probes under hybridization conditions, the presence of bands containing the probe sequences can be identified. The method is especially useful for identifying fragments in a restriction-enzyme DNA digest which contains a given probe sequence and for analyzing restriction-fragment length polymorphisms ("RFLPs").

Another approach to detecting the presence of a given sequence or sequences in a polynucleotide sample involves selective amplification of the sequence(s) by polymerase chain reaction. U.S. Patent No. 4,683,202 to Mullis, et al. and R.K. Saiki, et al., Science 230:1350 (1985). In this method, primers complementary to opposite end portions of the selected sequence(s) are used to promote, in conjunction with thermal cycling, successive rounds of primer-initiated replication. The amplified sequence(s) may be readily identified by a variety of techniques. This approach is particularly useful for detecting the presence of low-copy sequences in a polynucleotide-containing sample, e.g., for detecting pathogen sequences in a body-fluid sample.

More recently, methods of identifying known target sequences by probe ligation methods have been reported. U.S. Patent No. 4,883,750 to N.M. Whiteley, et al., D.Y. Wu, et al., Genomics 4:560 (1989). U. Landegren, et al., Science 241:1077 (1988), and E. Winn-Deen, et al., Clin. Chem. 37:1522 (1991). In one approach, known as oligonucleotide ligation assay ("OLA"), two probes or probe elements which span a target region of interest are hybridized to the target region. Where the probe elements basepair with adjacent target bases, the confronting ends of the probe elements can be joined by ligation, e.g., by treatment with ligase. The ligated probe element is then assayed, indicating the presence of the target sequence.

In a modification of this approach, the ligated probe elements act as a template for a pair of complementary probe elements. With continued cycles of denaturation, hybridization, and ligation in the presence of pairs of probe elements, the target sequence is amplified linearly, allowing very small amounts of target sequence to be detected and/or amplified. This approach is referred to as ligase detection reaction. When two complementary pairs of probe elements are utilized, the process is referred to as the ligase chain reaction which achieves exponential

amplification of target sequences. F. Barany, "Genetic Disease Detection and DNA Amplification Using Cloned Thermostable Ligase," Proc. Nat'l Acad. Sci. USA, 88:189-93 (1991) and F. Barany, "The Ligase Chain Reaction (LCR) in a PCR World," PCR Methods and Applications, 1:5-16 (1991).

Another scheme for multiplex detection of nucleic acid sequence differences is disclosed in U.S. Patent No. 5,470,705 to Grossman et. al. where sequence-specific probes, having a detectable label and a distinctive ratio of charge/translational frictional drag, can be hybridized to a target and ligated together. This technique was used in Grossman, et. al., "High-density Multiplex Detection of Nucleic Acid Sequences: Oligonucleotide Ligation Assay and Sequence-coded Separation," Nucl. Acids Res. 22(21):4527-34 (1994) for the large scale multiplex analysis of the cystic fibrosis transmembrane regulator gene.

Jou, et. al., "Deletion Detection in Dystrophia Gene by Multiplex Gap Ligase Chain Reaction and Immunochromatographic Strip Technology," Human Mutation 5:86-93 (1995) relates to the use of a so called "gap ligase chain reaction" process to amplify simultaneously selected regions of multiple exons with the amplified products being read on an immunochromatographic strip having antibodies specific to the different haptens on the probes for each exon.

There is a growing need (e.g., in the field of genetic screening) for methods useful in detecting the presence or absence of each of a large number of sequences in a target polynucleotide. For example, as many as 400 different mutations have been associated with cystic fibrosis. In screening for genetic predisposition to this disease, it is optimal to test all of the possible different gene sequence mutations in the subject's genomic DNA, in order to make a positive identification of "cystic fibrosis". It would be ideal to test for the presence or absence of all of the possible mutation sites in a single assay. However, the prior-art methods described above are not readily adaptable for use in detecting multiple selected sequences in a convenient, automated single-assay format.

Solid-phase hybridization assays require multiple liquid-handling steps, and some incubation and wash temperatures must be carefully controlled to keep the stringency needed for single-nucleotide mismatch discrimination. Multiplexing of this approach has proven difficult as optimal hybridization conditions vary greatly among probe sequences.

Developing a multiplex PCR process that yields equivalent amounts of each PCR product can be difficult and laborious. This is due to variations in the annealing rates of the primers in the reaction as well as varying polymerase extension rates for each sequence at a given Mg^{2+} concentration. Typically, primer, Mg^{2+}, and

salt concentrations, along with annealing temperatures are adjusted in an effort to balance primer annealing rates and polymerase extension rates in the reaction. Unfortunately, as each new primer set is added to the reaction, the number of potential amplicons and primer dimers which could form increases exponentially. Thus, with each added primer set, it becomes increasingly more difficult and time consuming to work out conditions that yield relatively equal amounts of each of the correct products.

Allele-specific PCR products generally have the same size, and an assay result is scored by the presence or absence of the product band(s) in the gel lane associated with each reaction tube. Gibbs et al., Nucleic Acids Res., 17:2437-48 (1989). This approach requires splitting the test sample among multiple reaction tubes with different primer combinations, multiplying assay cost. In PCR, discrimination of alleles can be achieved by attaching different fluorescent dyes to competing allelic primers in a single reaction tube (F.F. Chehab, et al., Proc. Natl. Acad. Sci. USA, 86:9178-9182 (1989)), but this route to multiplex analysis is limited in scale by the relatively few dyes which can be spectrally resolved in an economical manner with existing instrumentation and dye chemistry. The incorporation of bases modified with bulky side chains can be used to differentiate allelic PCR products by their electrophoretic mobility, but this method is limited by the successful incorporation of these modified bases by polymerase, and by the ability of electrophoresis to resolve relatively large PCR products which differ in size by only one of these groups. Livak et al., Nucleic Acids Res., 20:4831-4837 (1989). Each PCR product is used to look for only a single mutation, making multiplexing difficult.

Ligation of allele-specific probes generally has used solid-phase capture (U. Landegren et al., Science, 241:1077-1080 (1988); Nickerson et al., Proc. Natl. Acad. Sci. USA, 87:8923-8927 (1990)) or size-dependent separation (D.Y. Wu, et al., Genomics, 4:560-569 (1989) and F. Barany, Proc. Natl. Acad. Sci., 88:189-193 (1991)) to resolve the allelic signals, the latter method being limited in multiplex scale by the narrow size range of ligation probes. Further, in a multiplex format, the ligase detection reaction alone cannot make enough product to detect and quantify small amounts of target sequences. The gap ligase chain reaction process requires an additional step -- polymerase extension. The use of probes with distinctive ratios of charge/translational frictional drag for a more complex multiplex process will either require longer electrophoresis times or the use of an alternate form of detection.

The need thus remains for a rapid single assay format to detect the presence or absence of multiple selected sequences in a polynucleotide sample when

those sequences are in low abundance. Such detection is required when cancer-associated mutations are present in an excess of normal cells.

DNA Ligase

DNA ligases catalyze the formation of phosphodiester bonds at single-stranded breaks (nicks) in double-stranded DNA, and are required in DNA replication, repair, and recombination. The general mechanism of ligation reactions involves three reversible steps, as shown below for NAD^+-dependent ligases. In this example, the nicked DNA substrate is formed by annealing two short oligonucleotides (oligo **A** and **B**) to a longer complementary oligonucleotide. First, a covalently adenylated enzyme intermediate is formed by transfer of the adenylate group of NAD^+ to the e-NH_2 group of lysine in the enzyme. Second, the adenylate moiety is transferred from the enzyme to the 5'-terminal phosphate on oligo **B**. Finally, a phosphodiester bond is formed by a nucleophilic attack of the 3'-hydroxyl terminus of oligo **A** on the activated 5'-phosphoryl group of oligo **B** (Gumport, R.I., et al., Proc. Natl. Acad. Sci. USA, 68:2559-63 (1971); Modrich, P., et al., J. Biol. Chem., 248:7495-7501 (1973); Modrich, P., et al., J. Biol. Chem., 248:7502-11 (1973); Weiss, B., et al., J. Biol. Chem., 242:4270-72 (1967); Weiss, B., et al., J. Biol. Chem., 243:4556-63 (1968); Becker, A., et al., Proc. Natl. Acad. Sci. USA, 58:1996-2003 (1967); Yudelevich, A., et al., Proc. Natl. Acad. Sci. USA, 61:1129-36 (1968); Zimmerman, S.B., et al., Proc. Natl. Acad. Sci. USA, 57:1841-48 (1967); Zimmerman, S.B., et al., J. Biol. Chem., 244:4689-95 (1969); and Lehman, I.R., Science, 186:790-97 (1974)).

(i) $\text{E-(lys)-NH}_2 + \text{AMP~PRN}^+ \leftrightarrow \text{E-(lys)-NH}_2{}^+\text{~AMP} + \text{NMN}$

(ii) $\text{E-(lys)-NH}_2{}^+\text{~AMP} +$ 5' P-Oligo **B** $\leftrightarrow$ AMP~P-Oligo **B** $+ \text{E-(lys)-NH}_2$

(iii) Oligo **A**-3'OH + AMP~P-Oligo **B** $\leftrightarrow$ Oligo **A**-P-Oligo **B** + AMP

Within the last decade, genes encoding ATP-dependent DNA ligases have been cloned and sequenced from bacteriophages T3, T4, and T7 (Dunn, J.J., et al., J. Mol. Biol., 148:303-30 (1981); Armstrong, J., et al., Nucleic Acids Res., 11:7145-56 (1983); and Schmitt, M.P., et al., J. Mol. Biol., 193:479-95 (1987)), African swine fever virus (Hammond, J.M., et al., Nucleic Acids Res., 20:2667-71 (1992)), Vaccinia virus (Smith, G.L., et al., Nucleic Acids Res., 17:9051-62 (1989)), Shope fibroma virus (Parks, R.J., et al., Virology, 202:642-50 (1994)), an extremely thermophilic archaeon *Desulfurolobus ambivalens* (Kletzin, A., Nucleic Acids Res.,

20:5389-96 (1992)), *S. cerevisiae* (CDC9 gene)(Barker, D.G., et al., Mol. Gen. Genet., 200:458-62 (1985)); *S. pombe* (cdc17^{+}) (Barker, D.G., et al., Eur. J. Biochem., 162:659-67 (1988)), Xiphophorus (Walter, R.B., et al., Mol. Biol. Evol., 10:1227-38 (1993)); mouse fibroblast (Savini, E., et al., Gene, 144:253-57 (1994)); and *Homo sapiens* (human DNA ligase I, III, and IV) (Barnes, D.E., et al., Proc. Natl. Acad. Sci. USA, 87:6679-83 (1990); Chen, J., et al., Molec. and Cell. Biology, 15:5412-22 (1995); and Wei, Y.F., et al., Molec. & Cell. Biology, 15:3206-16 (1995)). In addition, five NAD^{+}-dependent bacterial DNA ligases have also been cloned: *E. coli* (Ishino, Y., et al., Mol. Gen. Genet., 204:1-7 (1986)), *Zymomonas mobilis* (Shark, K.B., et al., FEMS Microbiol. Lett., 96:19-26 (1992)), *Thermus thermophilus* (Barany, F., et al., Gene, 109:1-11 (1991) and Lauer, G., et al., J. Bacteriol., 173:5047-53 (1991)). *Rhodothermus marinus* (Thorbjarnardottir, S.H., et al., Gene, 161:1-6 (1995)), and *Thermus scotoductus* (Jonsson, Z.O., et al., Gene, 151:177-80 (1994)). ATP-dependent DNA ligases, as well as mammalian DNA ligases I and II contain a conserved active site motif, K(Y/A)DGXR, which includes the lysine residue that becomes adenylated (Tomkinson, A.E., et al, Proc. Natl. Acad. Sci. USA, 88:400-04 (1991) and Wang, Y.C., et al., J. Biol. Chem., 269:31923-28 (1994)). NAD^{+}-dependent bacterial DNA ligases contain a similar active site motif, KXDG, whose importance is confirmed in this work.

In vitro experiments using plasmid or synthetic oligonucleotide substrates reveal that T4 DNA ligase exhibits a relaxed specificity; sealing nicks with 3'- or 5'-AP sites (apurinic or apyrimidinic) (Goffin, C., et al., Nucleic Acids Res., 15(21):8755-71 (1987)), one-nucleotide gaps (Goffin, C., et al., Nucleic Acids Res., 15(21):8755-71 (1987)), 3'- and 5'-A–A or T-T mismatches (Wu, D.Y., et al., Gene, 76:245-54 (1989)), 5'-G-T mismatches (Harada, K., et al., Nucleic Acids Res., 21(10):2287-91 (1993)), 3'-C-A, C-T, T-G, T-T, T-C, A-C, G-G, or G-T mismatches (Landegren, U., et al., Science, 241:1077-80 (1988)). The apparent fidelity of T4 DNA ligase may be improved in the presence of spermidine, high salt, and very low ligase concentration, where only T-G or G-T mismatch ligations were detected (Wu, D.Y., et al., Gene, 76:245-54 (1989) and Landegren, U., et al., Science, 241:1077-80 (1988)). DNA ligase from *Saccharomyces cerevisiae* discriminates 3'-hydroxyl and 5'-phosphate termini separated by a one-nucleotide gap and 3'-A-G or T-G mismatches, however 5'-A-C, T-C, C-A, or G-A mismatches had very little effect on ligation efficiency (Tomkinson, A.E., et al., Biochemistry, 31:11762-71 (1992)). Mammalian DNA ligases I and III show different efficiencies in ligating 3' C-T, G-T, and T-G mismatches (Husain, I., et al., J. Biol. Chem., 270:9638-90 (1995)). The Vaccinia DNA virus efficiently discriminates against one-nucleotide, two-nucleotide

gaps and 3'-G-A, A-A, G-G, or A-G (purine-purine) mismatches, but easily seals 5'-C-T, G-T, T-T, A-C, T-C, C-C, G-G, T-G, or A-G mismatches as well as 3'-C-A, C-T, G-T, T-T, or T-G mismatches (Shuman, S., Biochem., 34:16138-47 (1995)).

The thermostable *Thermus thermophilus* DNA ligase (*Tth* DNA ligase) has been cloned and used in the ligase chain reaction (LCR) and ligase detection reaction (LDR) for detecting infectious agents and genetic diseases (Barany, F., Proc. Natl. Acad. Sci. USA, 88:189-93 (1991); Day, D., et al., Genomics, 29:152-62 (1995); Eggerding, F. A., PCR Methods and Applications, 4:337-45 (1995); Eggerding, F. A., et al., Human Mutation, 5:153-65 (1995); Feero, W., et al., Neurology, 43:668-73 (1993); Frenkel, L. M., et al., J. Clin. Micro., 33(2):342-47 (1995); Grossman, P. D., et al., Nucleic Acids Res., 22:4527-34 (1994); Iovannisci, D. M., et al., Mol. Cell. Probes, 7:35-43 (1993); Prchal, J. T., et al., Blood, 81:269-71 (1993); Ruiz-Opaz, N., et al., Hypertension, 24:260-70 (1994); Wiedmann, M., et al., Appl. Environ. Microbiol., 58:3443-47 (1992); Wiedmann, M., et al., Appl Environ Microbiol, 59(8):2743-5 (1993); Winn-Deen, E., et al., Amer. J. Human Genetics, 53:1512 (1993); Winn-Deen, E. S., et al., Clin. Chem., 40:1092 (1994); and Zebala, J., et al., "Detection of Leber's Hereditary Optic Neuropathy by nonradioactive-LCR. PCR Strategies," (Innis, M. A., Gelfand, D. H., and Sninsky, J. J., Eds.), Academic Press, San Diego (1996)). The success of these and future disease detection assays, such as identifying tumor associated mutations in an excess of normal DNA, depend on the exquisite fidelity of *Tth* DNA ligase.

Cancer Detection

As the second leading cause of death in this country, almost 600,000 people will die from cancer per year making cancer one of the most alarming of all medical diagnosis. Lifetime risks for developing invasive cancers in men and women are 50 percent and 33 percent, respectively. Expectations are that more than 1.2 million new cases of cancer will be diagnosed in the United States in 1995. Healthcare expenses for cancer in 1994 were approximately $104 billion. However, the full impact of cancer on families and society is not measured only by the amount of money spent on its diagnosis and treatment. A significant number of people are stricken with cancer in their most productive years. Cancers accounted for 18 percent of premature deaths in 1985 and in 1991 more than 9,200 women in the U.S. died from breast cancer before the age of 55. For colorectal and breast cancers, estimates are that nearly 140,000 and 183,000 new diagnoses, respectively, are predicted for 1995.

Currently, diagnosis of cancer is based on histological evaluation of tumor tissue by a pathologist. After a cancer is diagnosed, treatment is determined primarily by the extent or stage of the tumor. Tumor stage is defined by clinical, radiological, and laboratory methods. Standardized classification systems for the staging of tumors have been developed to clearly convey clinical information about cancer patients. Staging provides important prognostic information and forms the basis of clinical studies which allow the testing of new treatment strategies. A staging system was developed (TNM staging system), which classifies tumors according to the size of the primary tumor, the number of regional lymph nodes in which cancer is found, and the presence or absence of metastases to other parts of the body. Smaller cancers with no affected lymph nodes and no distant metastases are considered early stage cancers, which are often amenable to cure through surgical resection. A common measure of prognosis is the 5-year survival rate, the proportion of patients alive five years after the diagnosis of a cancer at a given stage. While 5-year survival rates for many cancers have improved over the last few decades, the fact that some early stage cancers recur within five years or later has led researchers to explore other additional prognostic markers including histological grade, cytometry results, hormone receptor status, and many other tumor markers. Most recently, investigators have explored the use of molecular alterations in cancers as markers of prognosis.

Genetic alterations found in cancers, such as point mutations and small deletions mentioned above, can act as markers of malignant cells.

Detection of Minority Nucleic Acid Sequences

A number of procedures have been disclosed to detect cancer using PCR. Sidransky, et. al., "Identification of *ras* Oncogene Mutations in the Stool of Patients with Curable Colorectal Tumors," Science 256: 102-05 (1992) detects colon cancer by identification of K-*ras* mutations. This involves a PCR amplification of total DNA, cloning into a phage vector, plating out the phage, repeated probing with individual oligonucleotides specific to several different K-*ras* mutations, and counting the percentage of positive plaques on a given plate. This is a technically difficult procedure which takes three days to complete, whereby the ratio of mutant to wild-type DNA in the stool sample is determined. Brennan, et. al., "Molecular Assessment of Histopathological Staging in Squamous-Cell Carcinoma of the Head and Neck," N. Engl. J. Med. 332(7): 429-35 (1995) finds p53 mutations by sequencing. This specific mutation is then probed for in margin tissue using PCR amplification of total DNA, cloning into a phage vector, plating out the phage, probing with an individual

oligonucleotide specific to the mutation found by sequencing, and counting the percentage of positive plaques on a given plate. Berthelemy, et. al., "Brief Communications--Identification of K-*ras* Mutations in Pancreatic Juice in the Early Diagnosis of Pancreatic Cancer," Ann. Int. Med. 123(3): 188-91 (1995) uses a PCR/restriction enzyme process to detect K-*ras* mutations in pancreatic secretions. This technique is deficient, however, in that mutations are not quantified. Similarly, Tada, et. al., "Detection of *ras* Gene Mutations in Pancreatic Juice and Peripheral Blood of Patients with Pancreatic Adenocarcinoma," Cancer Res. 53: 2472-74 (1993) and Tada, et. al., "Clinical Application of *ras* Gene Mutation for Diagnosis of Pancreatic Adenocarcinoma," Gastroent. 100: 233-38 (1991) subject such samples to allele-specific PCR to detect pancreatic cancer. This has the disadvantages of providing false positives due to polymerase extension off normal template, requiring electrophoretical separation of products to distinguish from primer dimers, being unable to multiplex closely-clustered sites due to interference of overlapping primers, being unable to detect single base or small insertions and deletions in small repeat sequences, and not being ideally suitable for quantification of mutant DNA in a high background of normal DNA. Hayashi, et. al., "Genetic Detection Identifies Occult Lymph Node Metastases Undetectable by the Histopathological Method," Cancer Res. 54: 3853-56 (1994) uses an allele-specific PCR technique to find K-*ras* or p53 mutations to identify occult lymph node metastases in colon cancers. A sensitivity of one tumor cell in one thousand of normal cells is claimed; however, obtaining quantitative values requires laborious cloning, plating, and probing procedures. In Mitsudomi, et. al., "Mutations of *ras* Genes Distinguish a Subset of Non-small-cell Lung Cancer Cell Lines from Small-cell Lung Cancer Cell Lines," Oncogene 6: 1353-62 (1991), human lung cancer cell lines are screened for point mutations of the K-, H-, and N-*ras* genes using restriction fragment length polymorphisms created through mismatched primers during PCR amplification of genomic DNA. The disadvantages of such primer-mediated RFLP include the requirement of electrophoretical separation to distinguish mutant from normal DNA, limited applicability to sites that may be converted into a restriction site, the requirement for additional analysis to determine the nature of the mutation, and the difficulty in quantifying mutant DNA in a high background of normal DNA. Further, these procedures tend to be laborious and inaccurate.

Coupled PCR/ligation processes have been used for detection of minority nucleotide sequences in the presence of majority nucleotide sequences. A PCR/LDR process is used in Frenkel, "Specific, Sensitive, and Rapid Assay for Human Immunodeficiency Virus Type 1 *pol* Mutations Associated with Resistance to

Zidovudine and Didanosine," J. Clin. Microbiol. 33(2): 342-47 (1995) to detect HIV mutants. This assay, however, cannot be used for multiplex detection. See also Abravaya, et. al., "Detection of Point Mutations With a Modified Ligase Chain (Gap-LCR)," Nucl. Acids Res. 23(4): 675-82 (1995) and Balles, et. al., "Facilitated Isolation of Rare Recombinants by Ligase Chain Reaction: Selection for Intragenic Crossover Events in the *Drosophila optomotor-blind* Gene," Molec. Gen. Genet. 245: 734-40 (1994).

Colorectal lesions have been detected by a process involving PCR amplification followed by an oligonucleotide ligation assay. See Jen, et. al., "Molecular Determinants of Dysplasia in Colorectal Lesions," Cancer Res. 54: 5523-26 (1994) and Redston, et. al., "Common Occurrence of *APC* and K-*ras* Gene Mutations in the Spectrum of Colitis-Associated Neoplasias," Gastroenter. 108: 383-92 (1995). This process was developed as an advance over Powell, et. al., "Molecular Diagnosis of Familial Adenomatous Polyposis," N. Engl. J. Med. 329(27): 1982-87 (1993). These techniques tend to be limited and difficult to carry out.

Other procedures have been developed to detect minority nucleotide sequences. Lu, et. al., "Quanititative Aspects of the Mutant Analysis by PCR and Restriction Enzyme Cleavage (MAPREC)" PCR Methods and Appl. 3: 176-80 (1993) detects virus revertants by PCR and restriction enzyme cleavage. The disadvantages of MAPREC include the requirement for electrophoretical separation to distinguish mutant from normal DNA, limited applicability to sites that may be converted into a restriction site, the requirement for additional analysis to determine the nature of the mutation, and difficulty in quantifying mutant DNA in a high background of normal DNA. In Kuppuswamy, et. al., "Single Nucleotide Primer Extension to Detect Genetic Diseases: Experimental Application to Hemophilia G (Factor IX) and Cystic Fibrosis Genes," Proc. Natl. Acad. Sci. USA 88: 1143-47 (1991), a PCR process is carried out using 2 reaction mixtures for each fragment to be amplified with one mixture containing a primer and a labeled nucleotide corresponding to the normal coding sequence, while the other mixture contains a primer and a labeled nucleotide corresponding to the mutant sequence. The disadvantages of such mini sequencing (i.e. SNuPe) are that the mutations must be known, it is not possible to multiplex closely clustered sites due to interference of overlapping primers, it is not possible to detect single base or small insertions and deletions in small repeat sequences, and four separate reactions are required. A mutagenically separated PCR process is disclosed in Rust, et. al., "Mutagenically Separated PCR (MS-PCR): a Highly Specific One Step Procedure for easy Mutation Detection" Nucl. Acids Res. 21(16): 3623-29 (1993) to distinguish normal and mutant alleles, using different length allele-specific

primers. The disadvantages of MS-PCR include possibly providing false positives due to polymerase extension off normal template, requiring electrophoretical separation of products to distinguish from primer dimers, the inability to multiplex closely-clustered sites due to interference of overlapping primers, the inability to detect single base or small insertions and deletions in small repeat sequences, and not being ideally suited for quantification of mutant DNA in high background of normal DNA. In Suzuki, et. al., "Detection of *ras* Gene Mutations in Human Lung Cancers by Single-Strand Conformation Polymorphism Analysis of Polymerase Chain Reaction Products," Oncogene 5: 1037-43 (1990), mutations are detected in a process having a PCR phase followed by phase involving single strand conformation polymorphism ("SSCP") of the amplified DNA fragments. The disadvantages of SSCP include the requirement for electrophoretical separation to distinghish mutant conformer from normal conformer, the failure to detect 30% of possible mutations, the requirement for additional analysis to determine the nature of the mutation, and the inability to distinguish mutant from silent polymorphisms.

Despite the existence of techniques for detecting minority nucleotide sequences in the presence of majority sequences, the need remains for improved procedures of doing so. It is particularly desirable to develop such techniques where minority nucleotide sequences can be quantified.

SUMMARY OF THE INVENTION

One aspect of the present invention relates to a method for detecting in a sample one or more minority target nucleotide sequences which differ from one or more majority target nucleotide sequences by one or more single-base changes, insertions, deletions, or translocations, wherein the minority target nucleotide sequences are present in the sample in lesser amounts than the majority nucleotide sequences.

One or more oligonucleotide probe sets are provided for use in conjunction with this method. Each set includes (a) a first oligonucleotide probe having a target-specific portion and (b) a second oligonucleotide probe having a target-specific portion. The oligonucleotide probes in a particular set are suitable for ligation together when hybridized adjacent to one another on a corresponding target nucleotide sequence, but have a mismatch which interferes with such ligation when hybridized to any other nucleotide sequence present in the sample.

The sample, the one or more oligonucleotide probe sets, and a ligase are blended to form a ligase detection reaction mixture. The ligase detection reaction

mixture is subjected to one or more ligase detection reaction cycles comprising a denaturation treatment and a hybridization treatment. In the denaturation treatment, any hybridized oligonucleotides are separated from the target nucleotide sequences. The hybridization treatment involves hybridizing the oligonucleotide probe sets at adjacent positions in a base-specific manner to the respective target nucleotide sequences, if present in the sample. The hybridized oligonucleotide probes from each set ligate to one another to form a ligation product sequence containing the target-specific portions connected together. The ligation product sequence for each set is distinguishable from other nucleic acids in the ligase detection reaction mixture. The oligonucleotide probe sets may hybridize to adjacent sequences in the sample other than the respective target nucleotide sequences but do not ligate together due to the presence of one or more mismatches. When hydridized oligonucleotide probes do not ligate, they individually separate during the denaturation treatment.

After the ligase detection reaction mixture is subjected to one or more ligase detection reaction cycles, ligation product sequences are detected. As a result, the presence of the minority target nucleotide sequence in the sample can be identified.

The second aspect of the present invention also relates to a method for identifying one or more of a plurality of sequences differing by one or more single-base changes, insertions, deletions, or translocations in a plurality of target nucleotide sequences. As noted above, a sample and one or more oligonucleotide probe sets are blended with a ligase to form a ligase detection reaction mixture. The ligase detection reaction mixture is subjected to one or more ligase detection reaction cycles, and the presence of ligation product sequences is detected. Here, however, a thermostable mutant ligase is utilized. This ligase is characterized by a fidelity ratio which is defined as the initial rate constant for ligating the first and second oligonucleotide probes hybridized to a target nucleotide sequence with a perfect match at the ligation junction between the target nucleotide sequence and the oligonucleotide probe having its 3' end at the ligation junction to the initial rate constant for ligating the first and second oligonucleotide probes hybridized to a target with a mismatch at the ligation junction between the target nucleotide sequence and the oligonucleotide probe having its 3' end at the ligation junction. The fidelity ratio for the thermostable mutant ligase is greater than the fidelity ratio for wild-type ligase.

The third aspect of the present invention also relates to a method for identifying one or more of a plurality of sequences differing by one or more single-base changes, insertions, deletions, or translocations in a plurality of target nucleotide sequences. As noted above, a sample and one or more oligonucleotide probe sets are

blended with a ligase to form a ligase detection reaction mixture. The ligase detection reaction mixture is subjected to one or more ligase detection reaction cycles, and the presence of ligation product sequences is then detected. Here, however, with regard to the oligonucleotide probe sets, the oligonucleotide probe which has its 3' end at the junction where ligation occurs has a modification. This modification differentially alters the ligation rate when the first and second oligonucleotide probes hybridize to a minority target nucleotide sequence in the sample with a perfect match at the ligation junction between the minority target nucleotide sequence and the oligonucleotide probe having its 3' end at the ligation junction compared to the ligation rate when the first and second oligonucleotide probes hybridize to the sample's majority target nucleotide sequence with a mismatch at the ligation junction between the majority target nucleotide sequence and the nucleotide probe having its 3' end at the ligation junction. Ligation with the modified oligonucleotide probe has a signal-to-noise ratio, of the ligation product sequence amounts for the minority and majority target nucleotide sequences to the amount of ligation product sequences produced from the same amount of majority target sequence alone, which is greater than the signal-to-noise ratio for ligation using an oligonucleotide probe lacking the modification.

In developing a procedure for detection of cancer-associated mutations or the presence of a minority target nucleotide sequence, it is necessary for the procedure to be capable of diagnosing cancer at an early stage. This requires that at least one clonal mutation be identified and accurately quantified from clinical samples. An ideal test of this type would rapidly screen up to hundreds of common mutations in multiple genes. It must accurately quantify less than 1% mutant DNA in the presence of normal DNA and correctly distinguish many closely clustered mutations in a multiplex format. For point mutations generally and, in small repeat sequences particularly, small insertions and deletions must be accurately identified. There must be internal controls against false-positive results and amenability to high throughput automation. The LDR process of the present invention is able to meet all of these objectives.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 is a flow diagram depicting a PCR/LDR process for detection of cancer-associated mutations by electrophoresis or capture on an addressable array where wild-type allele-specific oligonucleotide probes are excluded from the LDR phase to avoid overwhelming signal from minority mutant target and no marker is added to the LDR phase.

Figure 2 is a flow diagram depicting a PCR/LDR process for detection of cancer-associated mutations by electrophoresis or capture on an addressable array where wild-type allele-specific oligonucleotide probes are excluded from the LDR phase to avoid overwhelming signal from minority mutant target and a marker is added to the LDR phase.

Figure 3 is a flow diagram depicting a PCR/LDR process for detection of cancer-associated mutations by electrophoresis or capture on an addressable array where wild-type allele-specific oligonucleotide probes are utilized in the LDR phase at low levels and/or are modified to yield less ligation product corresponding to the majority target. This prevents the signal from minority mutant target from being overwhelmed.

Figure 4 is a schematic diagram depicting the PCR/LDR process of Figure 1 using electrophoresis to separate ligation products.

Figure 5 is a schematic diagram depicting the PCR/LDR process of Figure 2 using electrophoresis to separate ligation products.

Figure 6 is a schematic diagram depicting the PCR/LDR process of Figure 3 using electrophoresis to separate ligation products.

Figure 7 is a schematic diagram depicting the PCR/LDR process of Figure 1 using an addressable array.

Figure 8 is a schematic diagram depicting the PCR/LDR process of Figure 2 using an addressable array.

Figure 9 is a schematic diagram depicting the PCR/LDR process of Figure 3 using an addressable array.

Figure 10 relates to the construction of the *Themus thermophilus* DNA ligase mutants at amino acid residue 294 using site-specific mutagenesis.

Figure 11 shows the site-directed mutagenesis of possible active site regions of *Tth* ligase. The horizontal bar represents the full-length *Tth* DNA ligase protein with the arrow indicating the C-terminal end. Dark hatched bars represent regions with strong homology between *Tth* (i.e. *Thermus thermophilus*) DNA ligase and *E. coli* ligase, while the light hatched bars indicate regions with less homology. Amino acid substitutions produced by site-directed mutagenesis at K118, D120, K294, R337, G339, C412, C415, C428, and C433 are indicated. Amino acid residues which are identical among known NAD^+-dependent ligases are underlined.

Figures 12A-B show the primers for making the mutant ligases of Figure 11.

Figures 13A-C show a schematic representation of oligonucleotides used for ligation assays. Probe sequences were derived from human eukaryotic

protein synthesis initiation factor eIF-4E (Rychlik, W., et al., Proc. Natl. Acad. Sci. USA, 84:945-49 (1987), which is hereby incorporated by reference). This random eukaryotic DNA sequence was chosen to avoid any false signal arising from bacterial DNA contamination in partially purified mutant *Tth* DNA ligase preparations. The melting temperature of probes were predicted using the nearest neighbor thermodynamic method (Breslauer, K. J., et al., Proc. Natl. Acad. Sci. USA, 83:3746-3750 (1986), which is hereby incorporated by reference) (OLIGO 4.0 program, National Biosciences Inc., Plymouth, MN). Figure 13A and Figure 13B represent the formation of nicked DNA duplex using one of the template strands, ALg, as an example. Shown in Figure 13A, 4 different nicked DNA substrates are formed by annealing the common fluorescently labeled oligonucleotide, com5F, and one of the discriminating oligos (RP5'A, RP5'C, RP5'G, RP5'T) to the template strand, ALg. In the Figure 13B, 4 different nicked DNA substrates are formed by annealing the fluorescently labeled oligonucleotide, com3F, and one of the discriminating oligos (LP3'A, LP3'C, LP3'G, LP3'T) to the template strand, ALg. The full set of all 16 combinations of match and mismatch base pairing are thus formed by using ALg, GLg, TLg, and CLg (shown in Figure 13C) as the template strand, which vary at the underlined base. Products formed by ligation to the common fluorescently labeled probe can be discriminated by size on denaturing polyacrylamide gel due to the incorporation of different length of "A" tails.

Figure 14 shows the sequences for the probes used to make the oligonucleotides of Figure 13.

Figures 15A-E show the fidelity of nick closure by *Tth* DNA ligase at the 3'-side of the nick. The ligase substrate (nicked DNA duplex), shown in Figure 15A, is formed by annealing the discriminating oligonucleotide LP3'(A, C, G, or T) with a phosphorylated common oligonucleotide (3'-fluorescently labeled, com3F) on the template strand. The discriminating base "N" on the 3'-side of the discriminating oligonucleotide, and the "n" in a template strand were varied to give all 16 possible combinations of base-paring. "A_m" represents the "A" tail at the 5'-end of a discriminating oligonucleotide. Reactions were carried out in 40 µl mixture containing 1 mM NAD^+, 12.5 nM (500 fmoles) of nicked DNA duplex substrates and 0.125 nM (5 fmoles)*Tth* DNA ligase at 65 °C. Aliquots (5 µl) were removed at 0 hr, 2 hr, 4 hr, 6 hr, 8 hr, and 23 hr and separated on denaturing polyacrylamide gels. Data was analyzed using Genescan version 1.2 software. Results are plotted using Deltagraph Pro3 Software. Figures 15B-E represent results obtained with the same discriminating oligo, but with a different template strand. In panel 3'-A (Figure 15B), the discriminating oligonucleotide was LP3'A. A-T (◆), A-C (Δ), A-A (∇), and A-G

(O) represent DNA substrates containing TLg, CLg, ALg, and GLg as the template strand, respectively. In panel 3'-G (Figure 15D), the discriminating oligonucleotide was LP3'G. G-C (♦) , G-T (Δ), G-A (∇), and G-G (O) represent DNA substrates with CLg, TLg, ALg, and GLg as the template strand, respectively. In panel 3'-C (Figure 15C), LP3'C was the discriminating oligonucleotide. C-G (♦), C-A (Δ), C-T (∇), and C-C (O) indicate DNA substrates containing GLg, ALg, TLg, and CLg as the template strand, respectively. In panel 3'-T (Figure 15E), LP3'T was the discriminating oligonucleotide. T-A (♦), T-G (Δ), T-T (∇), and T-C (O) represent DNA substrates containing ALg, GLg, TLg, or CLg as the template strand, respectively.

Figures 16 A-E show the fidelity of nick closure by a thermostable DNA ligase at the 5'-side of the nick. Reaction conditions were the same as in Figure 15 except that different discriminating and common oligonucleotides were used. The discriminating base was on the 5'-side of the nick (phosphorylated oligonucleotides RP5'A, C, G, or T), while the common oligonucleotide was on the 3'-side of the nick, and was 5'-labeled with FAM (i.e. 6-carboxyfluorescein, fluorescent dye used in sequencing and mutation detection). See Figure 16A. In panel 5'-A (Figure 16B), the discriminating oligonucleotide was RP5'A. A-T (♦), A-C (Δ), A-A (∇), and A-G (O) represent DNA substrates containing TLg, CLg, ALg, and GLg as the template strand, respectively. In panel 5'-G (Figure 16D), the discriminating oligonucleotide was RP5'G. G-C (♦) , G-T (Δ), G-A (∇), and G-G (O) represent DNA substrate with CLg, TLg, ALg, and GLg as the template strand, respectively. In panel 5'-C (Figure 16C), RP5'C was the discriminating oligonucleotide. C-G (♦), C-A (Δ), C-T (∇), and C-C (O) indicate DNA substrates containing GLg, ALg, TLg, and CLg as the template strand, respectively. In panel 5'-T (Figure 16E), RP5'T was the discriminating oligonucleotide. T-A (♦), T-G (Δ), T-T (∇), and T-C (O) represent DNA substrates containing ALg, GLg, TLg, or CLg as the template strand, respectively.

Figures 17A-C are diagrams of oligonucleotide probes containing a base analogue and mismatch in the third position on the 3' side of the nick.

Figure 18 shows a table for improving the fidelity of *Thermus thermophilus* ligase. Sequences of probes LP3'C, LP3'T, and Com 3F are shown in Figure 14. Sequences for other discriminating probes are as the follows:

SLP3'C: 5' TACGTCTGCGGTGTTGCGTC 3'

SLP3'T: 5' CGTCTGCGGTGTTGCGTT 3'

SLP3'ATC: 5' ATGCGTCTGCGGTGTTGCATC 3'

SLP3'ATT: 5' GCGTCTGCGGTGTTGCATT 3'

SLP3'QTC: 5' AAATGCGTCTGCGGTGTTGCQTC 3'
SLP3'QTT: 5' ATGCGTCTGCGGTGTTGCQTT 3'

Base "N" in the discriminating oligonucleotide represents either C or T. "**Q**" indicates the Q base analogue. The template strand for all substrates tested except those containing Q base analogues is GLg, and its sequence is shown in Figure 14. The template strand in substrates containing Q base analogues is GLg.m3A which differs from GLg at a single site shown as bold. Initial rates of ligation (fmol/min) were calculated as the slope of the linear graph with the X-axis as the time in min., and the y-axis as the amount of products in fmol. Ligation fidelity of *Tth* DNA ligase is defined as a ratio of the initial rate of perfect match ligation over the initial rate of mismatch ligation.

Figure 19 shows the primers used for quantitative detection of single-base mutations in an excess of normal DNA by either wild-type or K294R *Tth* DNA ligase. LDR reactions contained 12.5 nM (250 fmole) of the mismatched template (GLg.m3A and GLg.m3A.rev = Normal template), containing from 0 to 2.5 nM (50 fmole) of perfect matched template (ALg.m3A and ALg.m3A.rev = Cancer template) in the presence of 25 fmol of either purified wild-type or mutant enzyme K294R. Each reaction was carried out in a 20 µl mixture containing 20 mM Tris-HCl, pH 7.6; 10 mM $MgCl_2$; 100 mM KCl; 10 mM DTT; 1 mM NAD^+; 25 nM (500 fmol) of the two short detecting primers and mixtures of templates. The reaction mixture was heated in GeneAmp 9600 (Perkin Elmer) for 15 sec. at 94°C before adding 25 fmol of the wild-type or mutant *Tth* DNA ligase. After incubation with the enzymes for another 30 sec at 94°C, LDR reactions were run for 15 sec at 94°C, and 4 min. at 65°C per cycle for 20 cycles. Reactions were completely stopped by chilling the tubes in an ethanol-dry ice bath, and adding 0.5 µl of 0.5 mM EDTA. Aliquots of 2.5 µl of the reaction products were mixed with 2.5 µl of loading buffer (83% Formamide, 8.3 mM EDTA, and 0.17% Blue Dextran) and 0.5 µl GeneScan Rox-1000 molecular weight marker, denatured at 94°C for 2 min, chilled rapidly on ice prior to loading on an 8 M urea-10% polyacrylamide gel, and electrophoresed on an ABI 373 DNA sequencer at 1400 volts. Fluorescent ligation products were analyzed and quantified using the ABI GeneScan 672 software.

Figures 20A-B show the LDR oligonucleotide probes (Figure 20A) and template sequences (Figure 20B) for a T:G mismatch in an excess of normal DNA by either wild-type or mutant K294R *Tth* DNA ligase. GLg.m3A and GLg.m3A.rev represent the mismatched template (Normal template), whereas ALg.m3A and ALg.m3A.rev represent the perfect matched template (Cancer

BNSDOCID: <WO 9803673A1 IA>

template); Primers SLP3'TTT represents the Normal primer for the perfect matched (Cancer) template; whereas SLP3'TTC represents the Normal primer for the mismatched (Normal) template. Similarly, experiments with the Q-analogue for a T:G mismatch use primers SLP3'Q_2TT and SLP3'Q_{18}TT as Normal primers for the matched (Cancer) template.

Figures 21A-B show the quantitative detection of single-base mutations (a T:G mismatch) in an excess of normal DNA by either wild-type or K294R *Tth* DNA ligase. The amount of LDR product formed when 0, 0.025 nM (0.5 fmol), 0.05 nM (1 fmol), 0.125 nM (2.5 fmol), 0.25 nM (5 fmol), 1.25 nM (25 fmol), and 2.5 nM (50 fmol) of "Cancer" template was used in combination with 12.5 nM (250 fmol) of the "Normal" template. The reaction was carried out in the presence of 25 nM (500 fmol) of regular primers (SLP3'TTT and Com 610 3'F), and 1.25 nM (25 fmol) of the wild-type or K294R mutant enzymes. The oligonucleotide probes used in this reaction create a T:G mismatch on the "Normal" template and a T:A match on the "Cancer" template at the ligation junction. LDR reactions were run for 15 sec at 94°C, and 4 min. at 65°C per cycle for 20 cycles. Reactions were completely stopped by chilling the tubes in an ethanol-dry ice bath and adding 0.5 µl of 0.5 mM EDTA. Aliquots of 2.5 µl of the reaction products were mixed with 2.5 µl of loading buffer (83% Formamide, 8.3 mM EDTA, and 0.17% Blue Dextran) and 0.5 µl GeneScan Rox-1000 molecular weight marker, denatured at 94°C for 2 min, chilled rapidly on ice prior to loading on an 8 M urea-10% polyacrylamide gel, and electrophoresed on an ABI 373 DNA sequencer at 1400 volts. Fluorescent ligation products were analyzed and quantified using the ABI GeneScan 672 software. The table (Figure 21B) describes the raw data for graph (Figure 21A). The data were analyzed and parameters of an exponential equation were fit to the data using the Deltagraph Pro 3.5. software. The X-axis indicated the different amounts of Cancer Template in 12.5 nM (250 fmol) of the Normal Template, while the Y-axis indicated the amount of LDR product generated. (■) represents 1.25 nM (25 fmol) of the wild-type enzyme whereas (●) represents 1.25 nM (25 fmol) of mutant K294 R enzyme.

Figures 22A-B show the signal-to-noise ratio of the LDR product with different concentrations of mutant template in normal DNA. The single-base mutation template ("Cancer") was diluted into 12.5 nM (250 fmol) of the "Normal" template, and assayed using 0.125 nM (25 fmol) of either wild-type or mutant K294R *Tth* DNA ligase. The oligonucleotide probes used in this reaction create a T:G mismatch on the "Normal" template and a T:A match on the "Cancer" template at the ligation junction. The signal-to-noise ratio is defined as the ratio of the amount of product formed with "Cancer" Template in the presence of "Normal" template (12.5

nM = 250 fmol template) to the amount of product formed by the same amount of Normal template alone. The table (Figure 22B) describes the raw data for the graph (Figure 22A). The data were analyzed and parameters of an exponential equation fit to the data using the Deltagraph Pro 3.5. software. The X-axis displayed different amounts of "Cancer" template in 12.5 nM (250 fmol) of the "Normal" template, while the Y-axis indicated the amount of LDR product generated. (■) represents 1.25 nM (25 fmol) of the wild-type enzyme whereas (●) represents 1.25 nM (25 fmol) of mutant K294 R enzyme.

Figures 23A-B show the amount of LDR product formed when 0, 0.005 nM (0.1 fmol), 0.0125 nM (0.25 fmol), 0.025 nM (0.5 fmol), 0.05 nM (1.0 fmol), 0.125 nM (2.5 fmol), 0.25 nM (5 fmol), and 0.5 nM (10 fmol) of Normal template (Glg.m3A and GLg.m3.Arev), are mixed with 25 nM (500 fmol) of regular primers (SLP3'TTC and Com 610 3'F), and 1.25 nM (25 fmol) of the wild-type or mutant enzymes. The primers used in this reaction create a C:G match on the "Normal" template at the ligation junction. LDR reactions were run for 15 sec at 94°C, and 4 min. at 65°C per cycle for 20 cycles. Reactions were completely stopped by chilling the tubes in an ethanol-dry ice bath, and adding 0.5 µl of 0.5 mM EDTA. Aliquots of 2.5 µl of the reaction products were mixed with 2.5 µl of loading buffer (83% Formamide, 8.3 mM EDTA, and 0.17% Blue Dextran) and 0.5 µl GeneScan Rox-1000 molecular weight marker, denatured at 94°C for 2 min, chilled rapidly on ice prior to loading on an 8 M urea-10% polyacrylamide gel, and electrophoresed on an ABI 373 DNA sequencer at 1400 volts. Fluorescent ligation products were analyzed and quantified using the ABI GeneScan 672 software. The table (Figure 23B) describes the raw data for the graph (Figure 23A). The data were analyzed and parameters of an exponential equation fit to the data using the Deltagraph Pro 3.5. software. The X-axis indicated the different amounts of Cancer Template in 12.5 nM (250 fmol) of the Normal Template, while the Y-axis indicated the amount of LDR product generated. (■) represents 1.25 nM (25 fmol) of the wild-type enzyme whereas (●) represents 1.25 nM (25 fmol) of mutant K294 R enzyme.

Figure 24 shows nucleotide analogue containing primers used for assaying ligase fidelity. Four different conditions were used to assess the fidelity of the wild-type and mutant *Tth* DNA ligase in a typical LDR assay. Each reaction was carried out in a 20 µl mixture containing 20 mM Tris-HCl, pH 7.6; 10 mM $MgCl_2$; 100 mM KCl; 10 mM DTT; 1 mM NAD^+; 25 nM (500 fmol) of the two short detecting primers and 12.5 nM (250 fmol) of the normal template when used alone, or 125 nM (2.5 fmol), and 0.5 nM (10 fmol), of the cancer template when used together with the normal template in a ratio of 1:100 and 1:25, respectively. The

oligonucleotide probes used in this reaction create a T:G mismatch on the "Normal" template and an T:A match on the "Cancer" template at the ligation junction. In addition, oligonucleotide probes SLP3'QTT create a Q_2:A or Q_{18}:T pairing at the 3rd position from the 3' end. The reaction mixture was heated in a GeneAmp 9600 thermocycler (Perkin Elmer) for 1.5 sec. at 94°C before adding 25 fmol of the wild-type and mutant *Tth* DNA ligase. After incubation with the enzymes for another 30 sec, LDR reactions were run for 15 sec at 94°C, and 4 min. at 65°C per cycle for 20 cycles. Reactions were completely stopped by chilling the tubes in an ethanol-dry ice bath, and adding 0.5 µl of 0.5 mM EDTA. 2.5 µl of reaction product was mixed with 2.5 µl of loading buffer (83% Formamide, 8.3 mM EDTA, and 0.17% Blue Dextran) and 0.5 µl Gene Scan Rox-1000 molecular weight marker, denatured at 94°C for 2 min, chilled rapidly on ice prior to loading on an 8 M urea-10% polyacrylamide gel, and electrophoresed on an ABI 373 DNA sequencer. Fluorescent labeled ligation products were analyzed and quantified using the ABI Gene Scan 672 software.

Figures 25A-D show different forms of oligonucleotide probes with nucleotide analogues for the LDR phase of the PCR/LDR process of the present invention.

Figure 26 shows primers used for quantitative detection of single-base mutations in an excess of normal DNA by either wild-type or K294R *Tth* DNA ligase (C:A mismatch). LDR reactions contained 12.5 nM (250 fmole) of the mismatched template (ALg.m3A and ALg.m3A.rev = Normal template) containing from 0 to 2.5 nM (50 fmole) of perfect matched template (GLg.m3A and GLg.m3A.rev = Cancer template), in the presence of 25 fmol of either purified wild-type or mutant enzyme K294R. Each reaction was carried out in a 20 µl mixture containing 20 mM Tris-HCl, pH 7.6; 10 mM $MgCl_2$; 100 mM KCl; 10 mM DTT; 1 mM NAD^+; 25 nM (500 fmol) of the two short detecting oligonucleotides probes and mixtures of templates. The probes used in this reaction create a C:A mismatch on the "Normal" template and an C:G match on the "Cancer" template at the ligation junction. The reaction mixture was heated in GeneAmp 9600 (Perkin Elmer) for 1.5 sec. at 94°C before adding 25 fmol of the wild-type or mutant *Tth* DNA ligase. After incubation with the enzymes for another 30 sec, LDR reactions were run for 15 sec at 94°C, and 4 min. at 65°C per cycle for 20 cycles. Reactions were completely stopped by chilling the tubes in an ethanol-dry ice bath, and adding 0.5 µl of 0.5 mM EDTA. Aliquots of 2.5 µl of the reaction products were mixed with 2.5 µl of loading buffer (83% Formamide, 8.3 mM EDTA, and 0.17% Blue Dextran) and 0.5 µl GeneScan Rox-1000 molecular weight marker, denatured at 94°C for 2 min, chilled rapidly on ice prior to loading on an 8 M urea-10% polyacrylamide gel, and electrophoresed on an ABI 373 DNA sequencer at

1400 volts. Fluorescent ligation products were analyzed and quantified using the ABI GeneScan 672 software.

Figures 27A-B show the quantitative detection of single-base mutations (a C:A mismatch) in an excess of normal DNA by either wild-type or K294R *Tth* DNA ligase. The amount of LDR product formed when 0, 0.025 nM (0.5 fmol), 0.05 nM (1 fmol), 0.125 nM (2.5 fmol), 0.25 nM (5 fmol), 1.25 nM (25 fmol), and 2.5 nM (50 fmol) of "Cancer" (i.e. Glg.m3A/Glg.m3A.rev) template was used in combination with 12.5 nM (250 fmol) of the "Normal" template (i.e. ALg.m3A/ALg.m3A.rev). The reaction was carried out in the presence of 25 nM (500 fmol) of regular oligonucleotide probes (SLP3'TTC and Com 610 3'F), and 1.25 nM (25 fmol) of the wild-type or mutant enzymes. The oligonucleotide probes used in this reaction create a C:A mismatch on the "Normal" template and a C:G match on the "Cancer" template at the ligation junction. LDR reactions were run for 15 sec at 94°C and 4 min. at 65°C per cycle for 20 cycles. Reactions were completely stopped by chilling the tubes in an ethanol-dry ice bath, and adding 0.5 µl of 0.5 mM EDTA. Aliquots of 2.5 µl of the reaction products were mixed with 2.5 µl of loading buffer (83% Formamide, 8.3 mM EDTA, and 0.17% Blue Dextran) and 0.5 µl GeneScan Rox-1000 molecular weight marker, denatured at 94°C for 2 min, chilled rapidly on ice prior to loading on an 8 M urea-10% polyacrylamide gel, and electrophoresed on an ABI 373 DNA sequencer. Fluorescent ligation products were analyzed and quantified using the ABI GeneScan 672 software. The table (Figure 26B) describes the raw data for the graph (Figure 26A). The data were analyzed and parameters of an exponential equation fit to the data using the Deltagraph Pro 3.5. software. The X-axis indicated the different amounts of Cancer Template in 12.5 nM (250 fmol) of the Normal Template, while the Y-axis indicated the amount of LDR product generated. (■) represents 1.25 nM (25 fmol) of the wild-type enzyme whereas (●) represents 1.25 nM (25 fmol) of mutant K294 R enzyme.

Figures 28A-B show the signal-to-noise ratio of the LDR product with different concentrations of the single-base mutation template ("Cancer") in combination with 12.5 nM (250 fmol) of the "Normal" template with 0.125 nM (25 fmol) by either wild-type or mutant K294R *Tth* DNA ligase. The signal-to-noise ratio is described as the ratio of the amount of product formed with varying concentrations of Cancer Template in the presence of 12.5 nM (250 fmol) of Normal template to the amount of product formed by 12.5 nM (250 fmol) of Normal template alone. The oligonucleotide probes (SLp3' TTC and Com 6103' F) used in this reaction create a C:A mismatch on the "Normal" template (ALg.m3A/ALg.m3A.rev) and an C:G match on the "Cancer" template (GLg.m3A/GLg.m3A.rev) at the ligation junction.

BNSDOCID: <WO 9803673A1 IA>

The table (Figure 28B) describes the raw data for the graph (Figure 28A). The data were analyzed and parameters of an exponential equation fit to the data using the Deltagraph Pro 3.5. software. The X-axis displayed different amounts of "Cancer" template in 12.5 nM (250 fmol) of the "Normal" template, while the Y-axis indicated the amount of LDR product generated. (■) represents 1.25 nM (25 fmol) of the wild-type enzyme whereas (●) represents 1.25 nM (25 fmol) of mutant K294 R enzyme.

Figures 29A-B show the amount of LDR product formed when 0, 0.005 nM (0.1 fmol), 0.0125 nM (0.25 fmol), 0.025 nM (0.5 fmol), 0.05 nM (1.0 fmol), 0.125 nM (2.5 fmol), 0.25 nM (5 fmol), and 0.5 nM (10 fmol) of Normal template (i.e. ALg.m3A/ALg.m3A.rev) alone, are reacted with 25 nM (500 fmol) of regular primers (SLP3'TTT and Com 610 3'F), and 1.25 nM (25 fmol) of the wild-type or mutant enzymes. LDR reactions were run for 15 sec at 94°C and 4 min. at 65°C per cycle for 20 cycles. The oligonucleotide probes used in this reaction create a T:A match on the "Normal" template at the ligation junction. Reactions were completely stopped by chilling the tubes in an ethanol-dry ice bath, and adding 0.5 µl of 0.5 mM EDTA. Aliquots of 2.5 µl of the reaction products were mixed with 2.5 µl of loading buffer (83% Formamide, 8.3 mM EDTA, and 0.17% Blue Dextran) and 0.5 µl GeneScan Rox-1000 molecular weight marker, denatured at 94°C for 2 min, chilled rapidly on ice prior to loading on an 8 M urea-10% polyacrylamide gel, and electrophoresed on an ABI 373 DNA sequencer at 1400 volts. Fluorescent ligation products were analyzed and quantified using the ABI GeneScan 672 software. The table (Figure 29B) describes the raw data for the graph (Figure 29A). The data were analyzed and parameters of an exponential equation fit to the data using the Deltagraph Pro 3.5. software. The X-axis indicated the amounts of the Normal Template used, while the Y-axis indicated the amount of LDR product generated. (■) represents 1.25 nM (25 fmol) of the wild-type enzyme whereas (●) represents 1.25 nM (25 fmol) of mutant K294 R enzyme.

Figures 30A-C show a scheme for PCR/LDR detection of mutations in codons 12, 13, and 61 of K-*ras*. At the top of the drawing (Figure 30A) is a schematic representation of the chromosomal DNA containing the K-*ras* gene. Exons are shaded and the position of codons 12, 13, and 61 shown. Exon-specific primers are used to amplify selectively K-*ras* DNA flanking these three codons. The middle (Figure 30B) and bottom (Figure 30C) of the diagram gives a schematic representation of primer design for LDR detection of all possible amino acid changes in codons 12, 13, and 61. For example, codon 12 (GGT) may mutate to GAT, GCT, or GTT. Allele-specific LDR oligonucleotide probes contain the discriminating base on the 3' end and a fluorescent label on the 5' end. Common oligonucleotides are

phosphorylated on the 5' end and contain a poly-A tail and blocking group on the 3' end. Different mutations are distinguished by separating the products on a polyacrylamide gel. Note that LDR oligonucleotide probes used for detecting mutations at codon 12 may interfere with hybridization of oligonucleotide probes used to detect mutations at codon 13. It will be necessary to determine experimentally if these probes can correctly identify mutant signal in the presence of the other LDR probes.

Figures 31A-B provide the name and sequence of K-*ras* LDR oligonucleotide probes used to detect different mutations at codons 12, 13, and 61 in a typical LDR reaction in an excess of normal DNA by either wild-type or mutant K294R *Tth* DNA ligase.

Figures 32 and 33A-B show the quantitative detection of a Gly->Asp mutation (C:A mismatch) in Codon 12 of the K-*ras* gene (G12D) in an excess of normal K-*ras* sequence by either wild-type or K294R *Tth* DNA ligase. The eleven lanes on the left of gel # mk960423 (#1-11) of Figure 32 represent data obtained when using wild-type *Tth* DNA ligase, while the eleven lanes on the right (#13-23) of Figure 32 represent data obtained when using the mutant *Tth* DNA ligase, K294R. The first three lanes in each case are negative controls without any added mutant K-*ras* sequence. The next eight lanes depict the amount of LDR product formed when 5 nM (100 fmol), 2 nM (40 fmol), 0.8 nM (20 fmol), 0.4 nM (8 fmol), 0.2 nM (4 fmol), 0.1 nM (2.0 fmol), 0.05 nM (1 fmol), and 0.025 nM (0.5 fmol) of mutant K-*ras* template, respectively, was used in combination with 100 nM (2000 fmol) of the wild-type K-*ras* DNA. The reaction was carried out in the presence of 25 nM (500 fmol) of one discriminating oligonucleotide probes (Fam-K-*ras* c12, 2D, and a common primer K-*ras* c12 Com-2) and 5 nM (100 fmol) of the wild-type or K294R mutant enzymes. The LDR probe for detecting the Gly->Asp mutation used in this reaction creates a C:A mismatch on the wild-type template and a T:A match on the Gly->Asp mutant K-*ras* template at the ligation junction. PCR reactions were run for 15 sec at 94°C, 1 min at 55°C, and 1 min. (+ 3 sec/cycle) at 72°C per cycle for 30 cycles. LDR reactions were run for 15 sec at 94°C, and 4 min. at 65°C per cycle for 20 cycles. Reactions were completely stopped by chilling the tubes in an ethanol-dry ice bath, and adding 0.5 µl of 0.5 mM EDTA. Aliquots of 2.5 µl of the reaction products were mixed with 2.5 µl of loading buffer (83% Formamide, 8.3 mM EDTA, and 0.17% Blue Dextran) and 0.5 µl GeneScan TAMRA 350 molecular weight marker, denatured at 94°C for 2 min, chilled rapidly on ice prior to loading on an 8 M urea-10% polyacrylamide gel, and electrophoresed on an ABI 373 DNA sequencer at 1400 volts. Fluorescent ligation products were quantified using the ABI GeneScan

672 software. The table (Figure 33B) describes the raw data for the graph (Figure 33A). The data were analyzed and parameters of an exponential equation fit to the data using the Deltagraph Pro 3.5. software. The X-axis indicated the different amounts of G12D template in 100 nM (2000 fmol) of the wild-type template, while the Y-axis indicated the amount of LDR product generated. (■) represents 5 nM (100 fmol) of the wild-type enzyme whereas (●) represents 5 nM (100 fmol) of mutant K294 R enzyme.

Figures 34A-B show the signal-to-noise ratio of the LDR product with different concentrations of the K-*ras* gene (from 0.025 nM [0.5 fmol] to 5 nM [100 fmol]) containing a single-base mutation (G12D) in combination with 100 nM (2000 fmol) of the wild-type K-*ras* template using 5 nM (100 fmol) of either wild-type or mutant K294R *Tth* DNA ligase. The probes used in this reaction create a C:A mismatch on the wild-type template and a T:A match on the G12D template at the ligation junction. The signal-to-noise ratio is defined as the ratio of the amount of product formed with G12D template in the presence of wild-type template (100 nM = 2000 fmol template) to the amount of product formed by the same amount of wild-type template alone. The table (Figure 34B) describes the raw data for the graph (Figure 34A). The data were analyzed and parameters of an exponential equation fit to the data using the Deltagraph Pro 3.5. software. The X-axis displayed different amounts of G12D template in 100 nM (2000 fmol) of wild-type template, while the Y-axis indicated the amount of LDR product generated. (■) represents 5 nM (100 fmol) of the wild-type enzyme whereas (●) represents 5 nM (100 fmol) of mutant K294 R enzyme.

Figures 35 and 36A-B show the quantitative detection of a Gly->Val mutation (a C:T mismatch) in Codon 12 of the K-*ras* gene (G12V) in an excess of normal K-*ras* sequence by either wild-type or K294R *Tth* DNA ligase. The eleven lanes on the left of gel # mk960405 (#1-11) (Figure 35) represent data obtained when using wild-type *Tth* DNA ligase, while the eleven lanes on the right (#13-23) (Figure 35) represent data obtained when using the mutant *Tth* DNA ligase, K294R. The first three lanes (Figure 35) in each case are negative controls without any added mutant K-*ras* sequence. The next eight lanes (Figure 35) depict the amount of LDR product formed when 0.1 nM (2.0 fmol), 0.2 nM (4 fmol), 0.4 nM (8 fmol), 0.8 nM (20 fmol), 2 nM (40 fmol), 4 nM (80 fmol), 5 nM (100 fmol), and 20 nM (200 fmol) of mutant K-*ras* template was used in combination with 100 nM (2000 fmol) of the wild-type K-*ras* template. The reaction was carried out in the presence of 25 nM (500 fmol) of six discriminating probes (Tet-K-*ras* c12.2A, Tet-K-*ras* c12.1S, Tet-K-*ras* c12.1C, Fam-K-*ras* c12. 1R, Fam-K-*ras* c12. 2D, Fam-K-*ras* c12. 2V); 75 nM (1500 fmol) of

two common probes (K-*ras* c12 Com-2 and K-*ras* c12 Com-1) and 5 nM (100 fmol) of the wild-type or K294R mutant enzymes. Tet is tetrachlorinated-6-carboxyfluorescein; fluorescent dye used in sequencing/mutation detection, while Com is an oligonucleotide probe using a 3' amino modified C3CPG column. This column is designed to produce a primary amine to the 3' terminus of a target oligonucleotide. The primary use of these 3' amino modifiers are for subsequent labelling as diagnostic probes and to generate an oligonucleotide resistant to the 3' exonuclease activity. This set of probes is capable of detecting the presence of all six single-base mutations in Codon 12 of the K-*ras* gene in a multiplex reaction. The LDR probe for detecting the Gly->Val mutation used in this reaction creates a C:T mismatch on the wild-type template and an A:T match on the Gly->Val mutant K-*ras* template at the ligation junction. PCR reactions were run for 15 sec at 94°C, 1 min at 55°C, and 1 min. (+ 3 sec/cycle) at 72°C per cycle for 30 cycles. LDR reactions were run for 15 sec at 94°C, and 4 min. at 65°C per cycle for 20 cycles. Reactions were completely stopped by chilling the tubes in an ethanol-dry ice bath, and adding 0.5 µl of 0.5 mM EDTA. Aliquots of 2.5 µl of the reaction products were mixed with 2.5 µl of loading buffer (83% Formamide, 8.3 mM EDTA, and 0.17% Blue Dextran) and 0.5 µl GeneScan TAMRA 350 molecular weight marker, denatured at 94°C for 2 min, chilled rapidly on ice prior to loading on an 8 M urea-10% polyacrylamide gel, and electrophoresed on an ABI 373 DNA sequencer at 1400 volts. Fluorescent ligation products were quantified using the ABI GeneScan 672 software. The table (Figure 36B) describes the raw data for the graph (Figure 36A). The data were analyzed and parameters of an exponential equation fit to the data using the Deltagraph Pro 3.5. software. The X-axis indicated the different amounts of G12V template in 100 nM (2000 fmol) of the wild-type template, while the Y-axis indicated the amount of LDR product generated. (■) represents 5 nM (100 fmol) of the wild-type enzyme whereas (●) represents 5 nM (100 fmol) of mutant K294 R enzyme.

Figures 37A-B show the signal-to-noise ratio of the LDR product in 26 primer multiplex reaction with different concentrations of the K-*ras* gene (from 0.1 nM [2 fmol] to 10 nM [200 fmol]) containing a single-base mutation (G12V) in combination with 100 nM (2000 fmol) of the wild-type K-*ras* template using 5 nM (100 fmol) of either wild-type or mutant K294R *Tth* DNA ligase. The G12V specific probes used in this reaction create a C:T mismatch on the wild-type template and an A:T match on the G12V template at the ligation junction. The greatest background noise in this multiplexed reaction was from probes designed to detect Q61R, representing a G:T mismatch, which was about 10-fold higher than from probes designed to detect G12D, i.e. representing a C:A mismatch. For consistency with

other assays, the signal-to-noise ratio in this multiplexed assay is defined as the ratio of the amount of product formed with G12V templates in the presence of wild-type template (100 nM = 2000 fmol template) to the amount of G12D LDR product formed by the same amount of wild-type template alone (representing a C:A mismatch). The table (Figure 37B) describes the raw data for the graph (Figure 37A). The data were analyzed and parameters of an exponential equation fit to the data using the Deltagraph Pro 3.5. software. The X-axis displayed different amounts of "Cancer" template in 100 nM (2000 fmol) of the "Normal" template, while the Y-axis indicated the amount of LDR product generated. (■) represents 5 nM (100 fmol) of the wild-type enzyme whereas (●) represents 5 nM (100 fmol) of mutant K294 R enzyme.

Figures 38 and 39A-B show the quantitative detection of a Gly->Val mutation (a C:T mismatch) in Codon 12 of the K-*ras* gene in an excess of normal K-*ras* sequence by either wild-type or K294R *Tth* DNA ligase. The eleven lanes on the left of gel # mk960429 (#1-11) (Figure 38) represent data obtained when using wild-type *Tth* DNA ligase, while the eleven lanes on the right (#13-23) (Figure 38) represent data obtained when using the mutant *Tth* DNA ligase, K294R. The first three lanes in each case are negative controls without any added mutant K-*ras* sequence. The next eight lanes depict the amount of LDR product formed when 0.1 nM (2.0 fmol), 0.2 nM (4 fmol), 0.4 nM (8 fmol), 0.8 nM (10 fmol), 2 nM (20 fmol), 4 nM (40 fmol), 5 nM (100 fmol), and 20 nM (200 fmol) of mutant K-*ras* template was used in combination with 100 nM (2000 fmol) of the wild-type K-*ras* template. The reaction was carried out in the presence of 25 nM (500 fmol) of nineteen discriminating primers (Tet-K-*ras* c12.2A, Tet-K-*ras* c12.1S, Tet-K-*ras* c12.1C, Tet-K-*ras* c13.4A, Tet-K-*ras* c13.3S, Tet-K-*ras* c13.3C, Tet-K-*ras* c61.7HT, Tet-K-*ras* c61.6R, Tet-K-*ras* c61.5K, Tet-K-*ras* c61.6P, Fam-K-*ras* c12.1R, Fam-K-*ras* c12.2D, Fam-K-*ras* c12.2V, Fam-K-*ras* c13.4D, Fam-K-*ras* c13.4V, Fam-K-*ras* c13.3R, Fam-K-*ras* c61.7HC, Fam-K-*ras* c61.6L, Fam-K-*ras* c61.5K); 50 nM (1000 fmol) of two common probes (K-*ras* c61 Com-7 and K-*ras* c12 Com-5); and 75 nM (1500 fmol) of five common primers (K-*ras* c12 Com-2, K-*ras* c12 Com-1, K-*ras* c13 Com-4, K-*ras* c13 Com-3, and K-*ras* c61 Com-6) and 5 nM (100 fmol) of the wild-type or K294R mutant enzymes. This set of probes is capable of detecting the presence of all nineteen mutations in Codons 12, 13, and 61 of the K-*ras* gene in a multiplex reaction. The LDR probe for detecting the Gly->Val mutation used in this reaction creates a C:T mismatch on the wild-type K-*ras* template and an A:T match on the Gly->Val mutant K-*ras* template at the ligation junction. PCR reactions were run for 15 sec at 94°C, 1 min at 55°C, and 1 min. (+ 3 sec cycle) at 72°C per cycle for 30

cycles. LDR reactions were run for 15 sec at 94°C, and 4 min. at 65°C per cycle for 20 cycles. Reactions were completely stopped by chilling the tubes in an ethanol-dry ice bath, and adding 0.5 µl of 0.5 mM EDTA. Aliquots of 2.5 µl of the reaction products were mixed with 2.5 µl of loading buffer (83% Formamide, 8.3 mM EDTA, and 0.17% Blue Dextran) and 0.5 µl GeneScan TAMRA 350 molecular weight marker, denatured at 94°C for 2 min, chilled rapidly on ice prior to loading on an 8 M urea-10% polyacrylamide gel, and electrophoresed on an ABI 373 DNA sequencer at 1400 volts. Fluorescent ligation products were quantified using the ABI GeneScan 672 software. The table (Figure 39B) describes the raw data for the graph (Figure 39A). The data were analyzed and parameters of an exponential equation fit to the data using the Deltagraph Pro 3.5. software. The X-axis indicated the different amounts of G12V template in 100 nM (2000 fmol) of the wild-type template, while the Y-axis indicated the amount of LDR product generated. (■) represents 5 nM (100 fmol) of the wild-type enzyme whereas (●) represents 5 nM (100 fmol) of mutant K294 R enzyme.

Figures 40A-B show the signal-to-noise ratio of the LDR product in a 26 primer multiplex reaction with different concentrations of the K-*ras* gene (from 0.1 nM [2 fmol] to 10 nM [200 fmol]) containing a single-base mutation (G12V) in combination with 100 nM (2000 fmol) of the wild-type K-*ras* template using 5 nM (100 fmol) of either wild-type or mutant K294R *Tth* DNA ligase. The G12V specific probes used in this reaction create a C:T mismatch on the wild-type template and an A:T match on the G12V template at the ligation junction. The greatest background noise in this multiplexed reaction was from probes designed to detect Q61R, representing a G:T mismatch, which was about 10-fold higher than from probes designed to detect G12D, i.e. representing a C:A mismatch. For consistency with our other assays, the signal-to-noise ratio in this multiplexed assay is defined as the ratio of the amount of product formed with G12V templates in the presence of wild-type template (100 nM = 2000 fmol template) to the amount of G12D LDR product formed by the same amount of wild-type template alone (representing a C:A mismatch). The table (Figure 40B) describes the raw data for the graph (Figure 40A). The data were analyzed and parameters of an exponential equation fit to the data using the Deltagraph Pro 3.5. software. The X-axis displayed different amounts of "Cancer" template in 100 nM (2000 fmol) of the "Normal" template, while the Y-axis indicated the amount of LDR product generated. (■) represents 5 nM (100 fmol) of the wild-type enzyme whereas (●) represents 5 nM (100 fmol) of mutant K294 R enzyme.

Figures 41-42 show the quantitative detection of mutations in the K-*ras* gene by K294 mutant *Tth* DNA ligase. The first lane (Figure 41) in gel # mk950513 is a negative control using wild-type K-*ras* DNA. The second lane (Figure 41) is a positive control which contains mutant Gly->Val K-*ras* DNA. The next twenty lanes (Figure 41) represent a blind test of LDR reactions on twenty samples containing different mutant K-*ras* DNA. The reactions were carried out in the presence of 25 nM (500 fmol) of nineteen discriminating primers (Tet-K-*ras* c12.2A, Tet-K-*ras* c12.1S, Tet-K-*ras* c12.1C, Tet-K-*ras* c13.4A, Tet-K-*ras* c13.3S, Tet-K-*ras* c13.3C, Tet-K-*ras* c61.7HT, Tet-K-*ras* c61.6R, Tet-K-*ras* c61.5K, Tet-K-*ras* c61.6P, Fam-K-*ras* c12.1R, Fam-K-*ras* c12.2D, Fam-K-*ras* c12.2V, Fam-K-*ras* c13.4D, Fam-K-*ras* c13.4V, Fam-K-*ras* c13.3R, Fam-K-*ras* c61.7HC, Fam-K-*ras* c61.6L, Fam-K-*ras* c61.5K); 50 nM (1000 fmol) of two common probes (K-*ras* c61 Com-7 and K-*ras* c12 Com-5); and 75 nM (1500 fmol) of five common probes (K-*ras* c12 Com-2, K-*ras* c12 Com-1, K-*ras* c13 Com-4, K-*ras* c13 Com-3, and K-*ras* c61 Com-6) and 5 nM (100 fmol) of the wild-type or K294R mutant enzymes. This set of probes is capable of detecting the presence of all nineteen mutations in Codons 12, 13, and 61, of the K-*ras* gene in a multiplex reaction. Microdissected tissue was transferred to a PCR tube, exposed to xylene for 10 min, washed 3 x in 95% ethanol, and desiccated. PCR reactions were run for 30 sec at 94°C, 1.5 min at 54°C, and 1 min. at 72°C per cycle for 35 cycles. LDR reactions were run for 15 sec at 94°C, and 4 min. at 65°C per cycle for 20 cycles. Reactions were completely stopped by chilling the tubes in an ethanol-dry ice bath, and adding 0.5 µl of 0.5 mM EDTA. Aliquots of 2.5 µl of the reaction products were mixed with 2.5 µl of loading buffer (83% Formamide, 8.3 mM EDTA, and 0.17% Blue Dextran) and 0.5 µl GeneScan TAMRA 350 molecular weight marker, denatured at 94°C for 2 min, chilled rapidly on ice prior to loading on an 8 M urea-10% polyacrylamide gel, and electrophoresed on an ABI 373 DNA sequencer at 1400 volts. Fluorescent ligation products were quantified using the ABI GeneScan 672 software. The data were analyzed and the results are presented in Figure 42 (called mutation) where they are compared with the results determined by dideoxy-sequencing (expected mutation).

Figure 43 is a table comparing 10 discordant samples from the PCR/LDR process described above with reference to Figures 41-42.

Figures 44 and 45A-B show the quantitative detection of different amounts of K-ras Normal template when varying amounts of wild-type probes were used by either wild-type or K294R *Tth* DNA ligase. Amount of LDR product formed when 25 nM (500 fmol), 50 nM (1000 fmol), and 100 nM (2000 fmol) of the "Normal" template was reacted with 0.5 nM (10 fmol), 2.5 nM (50 fmol), and 5 nM

(100 fmol) of the wild type discriminating probe (Tet-K-ras c12.2 WtG) and common probe (K-ras c12 Com-2) in the presence of 25 nM (500 fmol) of nineteen discriminating probes (Tet-K-ras c12.2A, Tet-K-ras c12.1S, Tet-K-ras c12.1C, Tet-K-ras c13.4A, Tet-K-ras c13.3S, Tet-K-ras c13.3C, Tet-K-ras c61.7HT, Tet-K-ras c61.6R, Tet-K-ras c61.5K, Tet-K-ras c61.6P, Fam-K-ras c12.1R, Fam-K-ras c12.2D, Fam-K-ras c 12.2V, Fam-K-ras c13.4D, Fam-K-ras c13.4V, Fam-K-ras c13.3R, Fam-K-ras c61.7HC, Fam-K-ras c61.6L, Fam-K-ras c61.5K); 50 nm (1000 fmol) of two common probes (K-ras c61 Com-7 and K-ras c12 Com-5); and 75 nm (1500 fmol) of five common probes (K-ras c12 Com-2, K-ras c12 Com-1, K-ras c13 Com-4, K-ras c13 Com-3, and K-ras c61 Com-6) and 5 nm (100 fmol) of the wild type or K294R mutant enzymes. LDR reactions were run for 15 sec at 94°C and 4 min. at 65°C per cycle for 20 cycles. The reactions were completely stopped by chilling the tubes in an ethanol-dry ice bath, and adding 0.5 µl of 0.5 mM EDTA. Aliquots of 2.5 µl of the reaction products were mixed with 2.5 µl of loading buffer (83% Formamide, 8.3 mM EDTA, and 0.17% Blue Dextran) and 0.5 µl GeneScan TAMRA 350 molecular weight marker, denatured at 94°C for 2 min, chilled rapidly on ice prior to loading on an 8 M urea-10% polyacrylamide gel, and electrophoresed on an ABI 373 DNA sequencer at 1400 volts. Fluorescent ligation products were analyzed and parameters of an exponential equation fit to the data using the Deltagraph Pro 3.5. software. The X-axis indicated the different amounts of the Normal Template, while the Y-axis indicated the amount of LDR product generated. (■, Δ ,□) represents 0.5 nM (10 fmol), 2.5 nM (50 fmol), and 5 nM (100 fmol), respectively of the wild type probes used with the wild-type enzyme whereas (●, ◆, ○) represents 0.5 (10 fmol), 2.5 nM (50 fmol), and 5 nM (100 fmol), of the wild type probes used with the K294R mutant enzyme.

DETAILED DESCRIPTION OF THE INVENTION

One aspect of the present invention relates to a method for detecting in a sample one or more minority target nucleotide sequences which differ from one or more majority target nucleotide sequences by one or more single-base changes, insertions, deletions, or translocations, wherein the minority target nucleotide sequences are present in the sample in lesser amounts than the majority nucleotide sequences.

One or more oligonucleotide probe sets are provided for use in conjunction with this method. Each set includes (a) a first oligonucleotide probe having a target-specific portion and (b) a second oligonucleotide probe having a

target-specific portion. The oligonucleotide probes in a particular set are suitable for ligation together when hybridized adjacent to one another on a corresponding target nucleotide sequence, but have a mismatch which interferes with such ligation when hybridized to any other nucleotide sequence present in the sample.

The sample, the one or more oligonucleotide probe sets, and a ligase are blended to form a ligase detection reaction mixture. The ligase detection reaction mixture is subjected to one or more ligase detection reaction cycles comprising a denaturation treatment and a hybridization treatment. In the denaturation treatment, any hybridized oligonucleotides are separated from the target nucleotide sequences. The hybridization treatment involves hybridizing the oligonucleotide probe sets at adjacent positions in a base-specific manner to the respective target nucleotide sequences, if present in the sample. The hybridized oligonucleotide probes from each set ligate to one another to form a ligation product sequence containing the target-specific portions connected together. The ligation product sequence for each set is distinguishable from other nucleic acids in the ligase detection reaction mixture. The oligonucleotide probe sets may hybridize to adjacent sequences in the sample other than the respective target nucleotide sequences but do not ligate together due to the presence of one or more mismatches. When hydridized oligonucleotide probes do not ligate, they individually separate during the denaturation treatment.

After the ligase detection reaction mixture is subjected to one or more ligase detection reaction cycles, ligation product sequences are detected. As a result, the presence of the minority target nucleotide sequence in the sample can be identified.

In effecting detection/quantification, there are 3 techniques of practicing the PCR/LDR process in accordance with the present invention with each being practiced using either of two formats. More particularly, the LDR phase can be carried out by (1) excluding wild-type allele-specific oligonucleotide probes from the LDR phase to avoid overwhelming signal from the minority mutant target and adding no marker, (2) excluding wild-type allele-specific oligonucleotide probes from the LDR phase but adding a marker to that phase, and (3) utilizing wild-type allele-specific oligonucleotide probes in the LDR phase at low levels and/or modified forms of those probes to yield less ligation product corresponding to the majority target which prevents signal from the minority mutant target from being overwhelmed. One detection format alternative involves use of capillary electrophoresis or gel electrophoresis and a fluorescent quantification procedure. Alternatively, detection can be carried out by capture on an array of capture oligonucleotide addresses and

fluorescent quantification. These alternatives are explained more fully with reference to Figures 1-9.

Figure 1 depicts the detection of cancer-associated mutations where wild-type allele-specific oligonucleotide probes are excluded from the LDR phase to avoid overwhelming signal from minority mutant target and no marker is added to the LDR phase. In step 1, after DNA sample preparation, multiple exons are subjected to PCR amplification using *Taq* polymerase under hot start conditions with target-specific oligonucleotide primers. The extension products produced during PCR are then diluted 1/20 during step 2. In step 3, the extension products are mixed with oligonucleotide probes containing allele-specific portions and common portions and the LDR phase of the process is initiated by addition of *Taq* ligase under hot start conditions. During LDR, oligonucleotide probes ligate to their adjacent oligonucleotide only in the presence of target sequence which gives perfect complementarity at the ligation junction. Absence of wild-type allele-specific oligonucleotide probes, and consequently absence of wild-type specific ligation product prevents the ligation detection reaction signal generated by minority mutant target from being overwhelmed.

The products can be detected by either of two formats. In the format of step 4a, products are separated by capillary gel electrophoresis, and fluorescent signals are quantified. On the other hand, in the format of step 4b, products are detected by specific hybridization to complementary sequences on an addressable array.

Figure 2 depicts the detection of cancer-associated mutations where wild-type allele-specific oligonucleotide probes are excluded from the LDR phase to avoid overwhelming signal from minority mutant target and a marker is added to the LDR phase. In step 1, after DNA sample preparation, multiple exons are subjected to PCR amplification using *Taq* polymerase under hot start conditions with target-specific oligonucleotide primers. Fluorescent quantification of PCR products can be achieved using capillary or gel electrophoresis in step 2. In step 3, the products are spiked with a 1/100 dilution of marker DNA (for each of the fragments). This DNA is homologous to wild type DNA, except it contains a mutation which is not observed in cancer cells, but which may be readily detected with the appropriate LDR probes. In step 4, the mixed DNA products from the PCR phase are then diluted 20-fold into fresh LDR buffer containing LDR oligonucleotide probes containing allele-specific portions and common portions. Step 5 involves the LDR phase of the process which is initiated by addition of *Taq* ligase under hot start conditions. During LDR, oligonucleotide probes ligate to their adjacent oligonucleotide probes only in

BNSDOCID: <WO___9803673A1_IA>

the presence of target sequence which gives perfect complementarity at the junction site.

The products may be detected in either of two formats. In the format of step 6a, products are separated by capillary or gel electrophoresis, and fluorescent signals are quantified. Ratios of mutant peaks to marker peaks give the approximate amount of cancer-associated mutations present in the original sample divided by 100. In the format of step 6b, products are detected by specific hybridization to complementary sequences on an addressable array. Ratios of fluorescent signals in mutant dots to marker dots give the approximate amount of cancer mutations present in the original sample divided by 100.

Figure 3 depicts the detection of additional cancer-associated mutations where additional wild-type allele-specific oligonucleotide probes are utilized in the LDR phase at low levels and/or are modified to yield less ligation product corresponding to the majority target. In step 1, after DNA sample preparation, multiple exons are subjected to PCR amplification using *Taq* polymerase under hot start conditions with target-specific oligonucleotide primers. The extension products produced during PCR are then diluted 1/20 during step 2. In step 3, the extension products are mixed with oligonucleotide probes containing allele-specific portions and common portions and the LDR phase of the process is initiated by addition of *Taq* ligase under hot start conditions. During LDR, oligonucleotide probes ligate to their adjacent oligonucleotide only in the presence of target sequence which gives perfect complementarity at the ligation junction. Due to the concentration and/or modification of the wild-type allele-specific oligonucleotide probes, the level of ligation product generated with these probes is comparable to the amount of ligation product generated from the minority target nucleotide sequences.

The products can be detected by either of two formats. In the format of step 4a, products are separated by capillary or gel electrophoresis, and fluorescent signal quantified. By way of example, consider the low level and/or modified wild-type allele-specific oligonucleotide probes ligating on a given amount of majority target nucleotide sequence (i.e. 1 picomole) generating the same amount of ligation product as generated from a given minority target sequence (using minority allele-specific oligonucleotide probes) present as a 100-fold dilution (i.e. 10 femtomoles) in the same amount (i.e. 1 picomole) of majority target nucleotide sequence. The ratio of mutant peaks to wild-type peaks gives the approximate amount of minority target (cancer-associated mutations) present in the original sample divided by 100. In the format of step 4b, products are detected by specific hybridization to complementary

sequences on an addressable array. Amount of minority product is quantified as described above.

The ligase detection reaction is described generally in WO 90/17239 to Barany et al., F. Barany et al., "Cloning, Overexpression and Nucleotide Sequence of a Thermostable DNA Ligase-Encoding Gene," Gene, 109:1-11 (1991), and F. Barany, "Genetic Disease Detection and DNA Amplification Using Cloned Thermostable Ligase," Proc. Natl. Acad. Sci. USA, 88:189-193 (1991), the disclosures of which are hereby incorporated by reference. In accordance with the present invention, the ligase detection reaction can use 2 sets of complementary oligonucleotides. This is known as the ligase chain reaction which is described in the 3 immediately preceding references, which are hereby incorporated by reference. Alternatively, the ligase detection reaction can involve a single cycle which is known as the oligonucleotide ligation assay. See Landegren, et al., "A Ligase-Mediated Gene Detection Technique," Science 241:1077-80 (1988); Landegren, et al., "DNA Diagnostics -- Molecular Techniques and Automation," Science 242:229-37 (1988); and U.S. Patent No. 4,988,617 to Landegren, et al., which are hereby incorporated by reference.

During the ligase detection reaction phase, the denaturation treatment is carried out at a temperature of 80-105 C, while hybridization takes place at 50-85 C. Each cycle comprises a denaturation treatment and a thermal hybridization treatment which in total is from about one to five minutes long. Typically, the ligation detection reaction involves repeatedly denaturing and hybridizing for 2 to 50 cycles. The total time for the ligase detection reaction phase is 1 to 250 minutes.

The oligonucleotide probe sets can be in the form of ribonucleotides, deoxynucleotides, modified ribonucleotides, modified deoxyribonucleotides, modified phosphate-sugar-backbone oligonucleotides (described *infra*), nucleotide analogs, and mixtures thereof.

In one variation, the oligonucleotides of the oligonucleotide probe sets each have a hybridization or melting temperature (i.e. T_m) of 66-70 C. These oligonucleotides are 20-28 nucleotides long.

The oligonucleotide probe sets, as noted above, have a reporter label suitable for detection. Useful labels include chromophores, fluorescent moieties, enzymes, antigens, heavy metals, magnetic probes, dyes, phosphorescent groups, radioactive materials, chemiluminescent moieties, and electrochemical detecting moieties.

The polymerase chain reaction process is fully described in H. Erlich, et. al., "Recent Advances in the Polymerase Chain Reaction," Science 252: 1643-50 (1991); M. Innis, et. al., PCR Protocols: A Guide to Methods and Applications,

Academic Press: New York (1990); and R. Saiki, et. al., "Primer-directed Enzymatic Amplification of DNA with a Thermostable DNA Polymerase," Science 239: 487-91 (1988), which are hereby incorporated by reference.

A particularly important aspect of the present invention is its capability to quantify the amount of target nucleotide sequence in a sample. This can be achieved in a number of ways by establishing standards which can be internal (i.e. where the standard establishing material is amplified and detected with the sample) or external (i.e. where the standard establishing material is not amplified, and is detected with the sample).

In accordance with one quantification method, the signal generated by ligation product sequences produced from the sample being analyzed, are detected. The strength of this signal is compared to a calibration curve produced from signals generated by ligation product sequences in samples with known amounts of target nucleotide sequence. As a result, the amount of target nucleotide sequence in the sample being analyzed can be determined. This technique involves use of an external standard.

Another quantification method, in accordance with the present invention, relates to an internal standard. Here, a known amount of one or more marker target nucleotide sequences is added to the sample. In addition, one or a plurality of marker-specific oligonucleotide probe sets are added along with the ligase, the previously-discussed oligonucleotide probe sets, and the sample to a mixture. The marker-specific oligonucleotide probe sets have (1) a first oligonucleotide probe with a target-specific portion complementary to the marker target nucleotide sequence, and (2) a second oligonucleotide probe with a target-specific portion complementary to the marker target nucleotide sequence and a detectable reporter label. The oligonucleotide probes in a particular marker-specific oligonucleotide set are suitable for ligation together when hybridized adjacent to one another on a corresponding marker target nucleotide sequence. However, there is a mismatch which interferes with such ligation when hybridized to any other nucleotide sequence present in the sample or added marker sequences. The presence of ligation product sequences is identified by detection of reporter labels. The amount of target nucleotide sequences in the sample is then determined by comparing the amount of ligation product sequence generated from known amounts of marker target nucleotide sequences with the amount of other ligation product sequences.

Another quantification method, in accordance with the present invention, involves analysis of a sample containing two or more of a plurality of target nucleotide sequences with a plurality of sequence differences. Here, ligation

product sequences corresponding to the target nucleotide sequences are detected and distinguished by any of the previously-discussed techniques. The relative amounts of the target nucleotide sequences in the sample are then quantified by comparing the relative amounts of ligation product sequences generated. This provides a quantitative measure of the relative level of the target nucleotide sequences in the sample.

Another quantification method, in accordance with the present invention, involves analysis of a sample containing two or more of a plurality of target nucleotide sequences with a plurality of sequence differences, where one or more target nucleotide sequences is in excess (majority) over other minority target nucleotide sequences. Here, in addition to the allele-specific oligonucleotide probes for the minority target nucleotide sequences, modified wild-type allele-specific oligonucleotide probes are also utilized in the LDR phase at low levels and/or are modified to yield less ligation product corresponding to the majority target. The presence of both minority target specific ligation products and majority target specific ligation products is identified by detection of reporter labels. The amount of minority target nucleotide sequences in the sample is determined by comparing the amount of low yield ligation product sequences generated from the majority target nucleotide sequences with the amount of other ligation products.

The preferred thermostable ligase is that derived from *Thermus aquaticus*. This enzyme can be isolated from that organism. M. Takahashi, et al., "Thermophillic DNA Ligase," J. Biol. Chem. 259:10041-47 (1984), which is hereby incorporated by reference. Alternatively, it can be prepared recombinantly. Procedures for such isolation as well as the recombinant production of *Thermus aquaticus* ligase (as well as *Thermus themophilus* ligase) are disclosed in WO 90/17239 to Barany, et. al., and F. Barany, et al., "Cloning, Overexpression and Nucleotide Sequence of a Thermostable DNA-Ligase Encoding Gene," Gene 109:1-11 (1991), which are hereby incorporated by reference. These references contain complete sequence information for this ligase as well as the encoding DNA. Other suitable ligases include *E. coli* ligase, T4 ligase, and *Pyrococcus* ligase.

The ligation detection reaction mixture may include a carrier DNA, such as salmon sperm DNA.

The hybridization step in the ligase detection reaction, which is preferably a thermal hybridization treatment discriminates between nucleotide sequences based on a distinguishing nucleotide at the ligation junctions. The difference between the target nucleotide sequences can be, for example, a single nucleic acid base difference, a nucleic acid deletion, a nucleic acid insertion, or

rearrangement. Such sequence differences involving more than one base can also be detected. Preferably, the oligonucleotide probe sets have substantially the same length so that they hybridize to target nucleotide sequences at substantially similar hybridization conditions. As a result, the process of the present invention is able to detect infectious diseases, genetic diseases, and cancer. It is also useful in environmental monitoring, forensics, and food science.

A wide variety of infectious diseases can be detected by the process of the present invention. Typically, these are caused by bacterial, viral, parasite, and fungal infectious agents. The resistance of various infectious agents to drugs can also be determined using the present invention.

Bacterial infectious agents which can be detected by the present invention include *Escherichia coli, Salmonella, Shigella, Klebsiella, Pseudomonas, Listeria monocytogenes, Mycobacterium tuberculosis, Mycobacterium avium-intracellulare, Yersinia, Francisella, Pasteurella, Brucella, Clostridia, Bordetella pertussis,* Bacteroides, *Staphylococcus aureus, Streptococcus pneumonia,* B-Hemolytic strep., *Corynebacteria, Legionella, Mycoplasma, Ureaplasma, Chlamydia, Neisseria gonorrhea, Neisseria meningitides, Hemophilus influenza, Enterococcus faecalis, Proteus vulgaris, Proteus mirabilis, Helicobacter pylori, Treponema palladium, Borrelia burgdorferi, Borrelia recurrentis, Rickettsial* pathogens, *Nocardia,* and *Acitnomycetes.*

Fungal infectious agents which can be detected by the present invention include *Cryptococcus neoformans, Blastomyces dermatitidis, Histoplasma capsulatum, Coccidioides immitis, Paracoccidioides brasiliensis, Candida albicans, Aspergillus fumigautus, Phycomycetes (Rhizopus), Sporothrix schenckii, Chromomycosis,* and *Maduromycosis.*

Viral infectious agents which can be detected by the present invention include human immunodeficiency virus, human T-cell lymphocytotrophic virus, hepatitis viruses (e.g., Hepatitis B Virus and Hepatitis C Virus), Epstein-Barr Virus, cytomegalovirus, human papillomaviruses, orthomyxo viruses, paramyxo viruses, adenoviruses, corona viruses, rhabdo viruses, polio viruses, toga viruses, bunya viruses, arena viruses, rubella viruses, and reo viruses.

Parasitic agents which can be detected by the present invention include *Plasmodium falciparum, Plasmodium malariae, Plasmodium vivax, Plasmodium ovale, Onchoverva volvulus, Leishmania, Trypanosoma* spp., *Schistosoma* spp., *Entamoeba histolytica, Cryptosporidium, Giardia* spp., *Trichimonas* spp., *Balatidium coli, Wuchereria bancrofti, Toxoplasma* spp., *Enterobius vermicularis, Ascaris*

lumbricoides, Trichuris trichiura, Dracunculus medinesis, trematodes, *Diphyllobothrium latum, Taenia* spp., *Pneumocystis carinii,* and *Necator americanis.*

The present invention is also useful for detection of drug resistance by infectious agents. For example, vancomycin-resistant *Enterococcus faecium,* methicillin-resistant *Staphylococcus aureus,* penicillin-resistant *Streptococcus pneumoniae,* multi-drug resistant *Mycobacterium tuberculosis,* and AZT-resistant human immunodeficiency virus can all be identified with the present invention.

Genetic diseases can also be detected by the process of the present invention. This can be carried out by prenatal or post-natal screening for chromosomal and genetic aberrations or for genetic diseases. Examples of detectable genetic diseases include: 21 hydroxylase deficiency, cystic fibrosis, Fragile X Syndrome, Turner Syndrome, Duchenne Muscular Dystrophy, Down Syndrome or other trisomies, heart disease, single gene diseases, HLA typing, phenylketonuria, sickle cell anemia, Tay-Sachs Disease, thalassemia, Klinefelter Syndrome, Huntington Disease, autoimmune diseases, lipidosis, obesity defects, hemophilia, inborn errors of metabolism, and diabetes.

Cancers which can be detected by the process of the present invention generally involve oncogenes, tumor suppressor genes, or genes involved in DNA amplification, replication, recombination, or repair. Examples of these include: BRCA1 gene, p53 gene, APC gene, Her2/Neu amplification, Bcr/Ab1, K-ras gene, and human papillomavirus Types 16 and 18. Various aspects of the present invention can be used to identify amplifications, large deletions as well as point mutations and small deletions/insertions of the above genes in the following common human cancers: leukemia, colon cancer, breast cancer, lung cancer, prostate cancer, brain tumors, central nervous system tumors, bladder tumors, melanomas, liver cancer, osteosarcoma and other bone cancers, testicular and ovarian carcinomas, head and neck tumors, and cervical neoplasms.

In the area of environmental monitoring, the present invention can be used for detection, identification, and monitoring of pathogenic and indigenous microorganisms in natural and engineered ecosystems and microcosms such as in municipal waste water purification systems and water reservoirs or in polluted areas undergoing bioremediation. It is also possible to detect plasmids containing genes that can metabolize xenobiotics, to monitor specific target microorganisms in population dynamic studies, or either to detect, identify, or monitor genetically modified microorganisms in the environment and in industrial plants.

The present invention can also be used in a variety of forensic areas, including for human identification for military personnel and criminal investigation,

paternity testing and family relation analysis, HLA compatibility typing, and screening blood, sperm, or transplantation organs for contamination.

In the food and feed industry, the present invention has a wide variety of applications. For example, it can be used for identification and characterization of production organisms such as yeast for production of beer, wine, cheese, yogurt, bread, etc. Another area of use is with regard to quality control and certification of products and processes (e.g., livestock, pasteurization, and meat processing) for contaminants. Other uses include the characterization of plants, bulbs, and seeds for breeding purposes, identification of the presence of plant-specific pathogens, and detection and identification of veterinary infections.

Desirably, the oligonucleotide probes are suitable for ligation together at a ligation junction when hybridized adjacent to one another on a corresponding target nucleotide sequence due to perfect complementarity at the ligation junction. However, when the oligonucleotide probes in the set are hybridized to any other nucleotide sequence present in the sample, there is a mismatch at a base at the ligation junction which interferes with ligation. Most preferably, the mismatch is at the base at the 3' base at the ligation junction. Alternatively, the mismatch can be at the bases adjacent to bases at the ligation junction.

As noted *supra*, detection and quantification can be carried out using capillary or gel electrophoresis or on a solid support with an array capture oligonucleotides.

Figure 4 is a schematic diagram depicting the PCR/LDR process of Figure 1 using electrophoresis to separate ligation products. More particularly, this diagram relates to detection of codon 12 of the K-ras gene which has a GGT sequence that codes for glycine ("Gly"). A small percentage of the cells contain the G to A mutation in GAT, which codes for aspartic acid ("Asp"). As illustrated, this process involves an initial PCR amplification in step 1, an LDR procedure in step 2, and a separation of fluorescent products followed by quantification in step 3. Alternatively, step 3 can involve ethidium bromide staining or running an additional LDR reaction on diluted product using normal oligonucleotide probes (See Figures 23 and 29 *infra*). The LDR probes for wild-type (i.e. normal) sequences are missing from the reaction. If the normal LDR probes (with the discriminating base being G) were included, they would ligate to the common probes and overwhelm any signal coming from the minority mutant target. Instead, as shown in Figure 4, the existence of a 44 base ligation product sequence with fluorescent label F1 coupled to a single nucleotide (designated N_1) and the A_n tail indicates the presence of the aspartic acid encoding mutant. This ligation product sequence has the same

F1 label as that formed by the existence of the arginine and valine encoded mutations. However, these sequences are distinguishable by virtue of their different lengths due to different length tails and different numbers of nucleotides N coupling the label to the remainder of the ligation product sequence. More particularly, the presence of an F1 labelled, 48 base ligation product sequence suggests the presence of the arginine encoding codon, while the presence of an F1 labelled, 46 base ligation product sequence indicates the presence of the valine encoding codon. These ligation product sequences are distinguished by size with the longer products having a lower electrophoretic mobility. The F2 labelled ligation products are similarly distinguished by their length which varies as a result of the different length tails and the number of nucleotides N coupling the label to the remainder of the ligation product sequence. More particularly, the 49 base ligation product sequence (due to 2 nucleotides N coupling the label and the A_{n+4} tail) indicates the presence of the cysteine encoding codon, the 47 base ligation product sequence (due to no nucleotides N coupling the label and the A_{n+4} tail) indicates the presence of the serine encoding codon, and the 45 base ligation product sequence (due to 2 nucleotides N coupling the label and the A_n tail) indicates the presence of the alanine encoding codon.

Figure 5 is a schematic diagram depicting the PCR/LDR process of Figure 2 using electrophoresis to separate ligation products. More particularly, this diagram relates to detection of codon 12 of the K-ras gene which has a GGT sequence that codes for glycine ("Gly"). A small percentage of the cells contain the G to A mutation in GAT, which codes for aspartic acid ("Asp"). As illustrated, this process involves an initial PCR amplification in step 1, an LDR procedure in step 2, and a separation of fluorescent products followed by quantification in step 3. After amplification, the PCR products are quantified. A marker template is added prior to the LDR phase where both allele-specific and marker-specific oligonucleotide probes are utilized. The LDR probes for wild-type (i.e. normal) sequences are missing from the reaction. If the normal LDR probes (with the discriminating base being G) were included, they would ligate to the common probes and overwhelm any signal coming from the minority mutant target. Instead, as shown in Figure 5, the existence of a 44 base ligation product sequence with fluorescent label F1 coupled to a single nucleotide (designated N_1) and the A_n tail indicates the presence of the aspartic acid encoding mutant. This ligation product sequence has the same F1 label as that formed by the existence of the arginine and valine encoded mutations. However, these sequences are distinguishable by virtue of their different lengths due to different length tails and different numbers of nucleotides N coupling the label to the remainder

of the ligation product sequence. More particularly, the presence of an F1 labelled, 48 base ligation product sequence suggests the presence of the arginine encoding codon, while the presence of an F1 labelled, 46 base ligation product sequence indicates the presence of the valine encoding codon. These ligation product sequences are distinguished by size with the longer products having a lower electrophoretic mobility. The F2 labelled ligation products are similarly distinguished by their length which varies as a result of the different length tails and the number of nucleotides N coupling the label to the remainder of the ligation product sequence. More particularly, the 49 base ligation product sequence (due to 2 nucleotides N coupling the label and the A_{n+4} tail) indicates the presence of the cysteine encoding codon, the 47 base ligation product sequence (due to no nucleotides N coupling the label and the A_{n+4} tail) indicates the presence of the serine encoding codon, and the 45 base ligation product sequence (due to 2 nucleotides N coupling the label and the A_n tail) indicates the presence of the alanine encoding codon. The ligation product formed by the marker-specific oligonucleotide probe is 43 bases and has the F_2 label (due to 0 nucleotides N coupling the label and the A_n tail). As discussed above, the amount of minority target nucleotide sequences in the sample is determined by comparing the amount of ligation product sequence generated from known amounts of marker target nucleotide sequences with the amount of other ligation product sequences.

Figure 6 is a schematic diagram depicting the PCR/LDR process of Figure 3 using electrophoresis to separate ligation products. More particularly, this diagram relates to detection of codon 12 of the K-ras gene which has a GGT sequence that codes for glycine ("Gly"). A small percentage of the cells contain the G to A mutation in GAT, which codes for aspartic acid ("Asp"). As illustrated, this process involves an initial PCR amplification in step 1, an LDR procedure in step 2, and a separation of fluorescent products followed by quantification in step 3. The LDR probes for wild-type (i.e. normal) sequences are used at low level and/or are modified to yield less ligation product sequence corresponding to wild type target nucleotide sequence. As shown in Figure 6, the existence of a 44 base ligation product sequence with fluorescent label F1 coupled to a single nucleotide (designated N_1) and the A_n tail indicates the presence of the aspartic acid encoding mutant. This ligation product sequence has the same F1 label as that formed by the existence of the arginine and valine encoded mutations. However, these sequences are distinguishable by virtue of their different lengths due to different length tails and different numbers of nucleotides N coupling the label to the remainder of the ligation product sequence. More particularly, the presence of an F1 labelled, 48 base ligation product sequence suggests the presence of the arginine encoding codon, while the presence of an F1

labelled, 46 base ligation product sequence indicates the presence of the valine encoding codon. These ligation product sequences are distinguished by size with the longer products having a lower electrophoretic mobility. The F2 labelled ligation products are similarly distinguished by their length which varies as a result of the different length tails and the number of nucleotides N coupling the label to the remainder of the ligation product sequence. More particularly, the 49 base ligation product sequence (due to 2 nucleotides N coupling the label and the A_{n+4} tail) indicates the presence of the cysteine encoding codon, the 47 base ligation product sequence (due to no nucleotides N coupling the label and the A_{n+4} tail) indicates the presence of the serine encoding codon, and the 45 base ligation product sequence (due to 2 nucleotides N coupling the label and the A_n tail) indicates the presence of the alanine encoding codon. The ligation product formed by the wild type allele-specific oligonucleotide probe is 43 bases and has the F2 label (due to 0 nucleotides N coupling the label and the A_n tail). In the labelled probe forming that ligation product, there is a base N located 3 base positions away from the ligation junction which can be either the conventional, proper nucleotide for the wild type target (if that probe is used at low level), or a mismatch, or a nucleotide base analogue. Use of a mismatched nucleotide, a nucleotide base analogue, and/or a modification in the sugar phosphate backbone reduces the amount of ligation product formed off wild-type target. Thus, the presence of wild type target can be detected without overwhelming the signal generated by the presence of minority mutant target. The amount of minority target nucleotide sequences in the sample is determined by comparing the amount of low yield ligation product sequences generated from the majority target nucleotide sequences with the amount of other ligation products.

Figures 4-6 show the use of the ligase detection reaction to detect mismatches at the 3' end of the distinguishing oligonucleotide probe. In other cases, however, the mismatch can be at the penultimate position to the 3' end or and the third position away from the 3' end.

The use of capillary and gel electrophoresis for such purposes is well known. See e.g., Grossman, et. al., "High-density Multiplex Detection of Nucleic Acid Sequences: Oligonucleotide Ligation Assay and Sequence-coded Separation," Nucl. Acids Res. 22(21): 4527-34 (1994), which is hereby incorporated by reference.

Figure 7 is a schematic diagram depicting the PCR/LDR process of Figure 1 for detection of cancer-associated mutations at adjacent alleles using an addressable array. Figure 7 relates to the detection of codon 12 of the K-ras gene which has a wild-type GGT sequence that codes for glycine ("Gly") and minority mutant GAT sequence coding for aspartic acid ("Asp"). The process of Figure 7

involves an initial PCR amplification in step 1, an LDR procedure in step 2, and capture on a solid support in step 3. As in Figure 4, the LDR probes for the wild-type target sequence are missing from the reaction to avoid overwhelming signal produced by the mutant target sequence. According to this embodiment of the present invention, as shown in Figure 7, the presence of the aspartic acid encoding GAT sequence produces a ligation product sequence with label F and addressable array-specific portion Z4. The existence of such a ligation product sequence is indicated by the presence of a nucleic acid, having label F, hybridized at an address on a solid support with a capture oligonucleotide complementary to addressable array-specific portion Z4. As shown in Figure 7, the support has an array of addresses with capture oligonucleotides complementary to different addressable array-specific portions Z1 to Z6. Since common oligonucleotide probes with label F are used, by observing which site on the solid support they hybridize to, different ligation product sequences are distinguished.

Figure 8 is a schematic diagram depicting the PCR/LDR process of Figure 2 for detection of cancer-associated mutations at adjacent alleles using an addressable array. Figure 8 relates to the detection of codon 12 of the K-ras gene which has a wild-type GGT sequence that codes for glycine ("Gly") and minority mutant GAT sequence coding for aspartic acid ("Asp"). The process of Figure 8 involves an inital PCR amplification in step 1, an LDR procedure in step 2, and capture on a solid support in step 3. As in Figure 5, the LDR probes for the wild-type target sequence are missing from the reaction to avoid overwhelming signal produced by the mutant target sequence. According to this embodiment of the present invention, as shown in Figure 8, the presence of the aspartic acid encoding GAT sequence produces a ligation product sequence with label F and addressable array-specific portion Z4. The existence of such a ligation product sequence is indicated by the presence of a nucleic acid, having label F, hybridized at an address on a solid support with a capture oligonucleotide complementary to addressable array-specific portion Z4. As shown in Figure 8, the support has an array of addresses with capture oligonucleotides complementary to different addressable array-specific portions Z1 to Z7. Since common oligonucleotide probes with label F are used, by observing which site on the solid support they hybridize to, different ligation product sequences are distinguished. The presence of ligation product sequence produced from a marker-specific probe is indicated by the existence of a nucleic acid, having label F, hybridized at an address on a solid support with a capture oligonucleotide complementary to addressable array-specific portion Z7. The amount of target nucleotide sequences in the sample is determined by comparing the amount of

ligation product sequence generated from known amounts of marker target nucleotide sequences with the amount of other ligation product sequences.

Figure 9 is a schematic diagram depicting the PCR/LDR process of Figure 3 for detection of cancer-associated mutations at adjacent alleles using an addressable array. Figure 9 relates to the detection of codon 12 of the K-ras gene which has a wild-type GGT sequence that codes for glycine ("Gly") and minority mutant GAT sequence coding for aspartic acid ("Asp"). The process of Figure 9 involves an inital PCR amplification in step 1, an LDR procedure in step 2, and capture on a solid support in step 3. The LDR probes for wild-type (i.e. normal) sequences are used at low level and/or are modified to yield less ligation product sequence corresponding to wild type target nucleotide sequence. According to this embodiment of the present invention, as shown in Figure 9, the presence of the aspartic acid encoding GAT sequence produces a ligation product sequence with label F and addressable array-specific portion Z4. The existence of such a ligation product sequence is indicated by the presence of a nucleic acid, having label F, hybridized at an address on a solid support with a capture oligonucleotide complementary to addressable array-specific portion Z4. As shown in Figure 9, the support has an array of addresses with capture oligonucleotides complementary to different addressable array-specific portions Z1 to Z7. Since common oligonucleotide probes with label F are used, different ligation product sequences are distinguished by which site on the solid support they hybridize to. The ligation product formed by the wild type allele-specific oligonucleotide probe is indicated by the existence of a nucleic acid, having label F, hybridized at an address on a solid support with a capture oligonucleotide complementary to addressable array-specific portion Z7. In the labelled probe forming that ligation product, there is a base N located 3 base positions away from the ligation junction that can be either a conventional nucleotide for the wild type target (if that probe is used at low level), or a mismatch nucleotide, or a nucleotide base analogue. Use of a mismatched nucleotide, a nucleotide base analogue, and/or a modification in the sugar phosphate backbone reduces the amount of ligation product formed off wild-type target. Thus, the presence of wild type target can be detected without overwhelming the signal generated by the presence of minority mutant target. The amount of minority target nucleotide sequences in the sample is determined by comparing the amount of low yield ligation product sequences generated from the majority target nucleotide sequences with the amount of other ligation products.

The use of a solid support with an array of capture oligonucleotides is fully disclosed in pending provisional U.S. Patent Application Serial No. 60/011,359, which is hereby incorporated by reference. When using such arrays, the

oligonucleotide probes used in the above-described LDR phase have an addressable array-specific portion. After the LDR phase is completed, the addressable array-specific portions for the products of such processes remain single stranded and are caused to hybridize to the capture oligonucleotides during a capture phase. See Newton, et al., "The Production of PCR Products With 5' Single-Stranded Tails Using Primers That Incorporate Novel Phosphoramidite Intermediates," Nucl. Acids Res. 21(5):1155-62 (1993), which is hereby incorporated by reference.

During the capture phase of the process, the mixture is contacted with the solid support at a temperature of 45-90 C and for a time period of up to 60 minutes. Hybridizations may be accelerated by adding cations, volume exclusion or chaotropic agents. When an array consists of dozens to hundreds of addresses, it is important that the correct ligation product sequences have an opportunity to hybridize to the appropriate address. This may be achieved by the thermal motion of oligonucleotides at the high temperatures used, by mechanical movement of the fluid in contact with the array surface, or by moving the oligonucleotides across the array by electric fields. After hybridization, the array is washed sequentially with a low stringency wash buffer and then a high stringency wash buffer.

It is important to select capture oligonucleotides and addressable nucleotide sequences which will hybridize in a stable fashion. This requires that the oligonucleotide sets and the capture oligonucleotides be configured so that the oligonucleotide sets hybridize to the target nucleotide sequences at a temperature less than that which the capture oligonucleotides hybridize to the addressable array-specific portions. Unless the oligonucleotides are designed in this fashion, false positive signals may result due to capture of adjacent unreacted oligonucleotides from the same oligonucleotide set which are hybridized to the target.

The capture oligonucleotides can be in the form of ribonucleotides, deoxyribonucleotides, modified ribonucleotides, modified deoxyribonucleotides, peptide nucleotide analogues, modified peptide nucleotide analogues, modified phosphate-sugar backbone oligonucleotides, nucleotide analogues, and mixtures thereof.

Where an array is utilized, the detection phase of the process involves scanning and identifying if LDR products have been produced and correlating the presence of such products to a presence or absence of the target nucleotide sequence in the test sample. Scanning can be carried out by scanning electron microscopy, confocal microscopy, charge-coupled device, scanning tunneling electron microscopy, infrared microscopy, atomic force microscopy, electrical conductance, and fluorescent or phosphor imaging. Correlating is carried out with a computer.

The present invention is useful in distinguishing a minority target nucleotide sequence from the majority nucleotide sequence in a sample at a respective ratio of 1:500 for a G:T or T:G mismatch between the majority target nucleotide sequence and one of the oligonucleotide probes. Further, this method can distinguish a minority target nucleotide sequence from the majority nucleotide sequence in a sample at a respective ratio of 1:2000 for other than a G:T or T:G mismatch between the majority target nucleotide sequence and one of the oligonucleotide probes.

For low abundance multiple allele differences at multiple nearby or adjacent positions, the process of the present invention distinguishes minority target nucleotide sequences from the majority target sequence at a respective ratio of 1:100 for all mismatches between the majority target nucleotide sequence and one of the oligonucleotide probes. In such situations, the minority target nucleotide sequence to majority target nucleotide sequence respective ratio is 1:500 for other than G:T or T:G mismatches between the majority target nucleotide sequence and one of the oligonucleotide probes.

The second aspect of the present invention also relates to a method for identifying one or more of a plurality of sequences differing by one or more single-base changes, insertions, deletions, or translocations in a plurality of target nucleotide sequences. As noted above, a sample and one or more oligonucleotide probe sets are blended with a ligase to form a ligase detection reaction mixture. The ligase detection reaction mixture is subjected to one or more ligase detection reaction cycles, and the presence of ligation product sequences is detected. Here, however, a thermostable mutant ligase is utilized. This ligase is characterized by a fidelity ratio which is defined as the initial rate constant for ligating the first and second oligonucleotide probes hybridized to a target nucleotide sequence with a perfect match at the ligation junction between the target nucleotide sequence and the oligonucleotide probe having its 3' end at the ligation junction to the initial rate constant for ligating the first and second oligonucleotide probes hybridized to a target with a mismatch at the ligation junction between the target nucleotide sequence and the oligonucleotide probe having its 3' end at the ligation junction. The fidelity ratio for the thermostable mutant ligase is greater than the fidelity ratio for wild-type ligase.

The use of a mutant ligase in accordance with the process of the present invention can be explained as follows. The specificity of an enzymatic reaction is determined by the catalytic constant, k_{cat}, and the apparent binding constant, K_M, and expressed as the specificity constant k_{cat}/K_M. Any modifications made on the enzyme itself, substrate, or reaction conditions, which affect k_{cat} or K_M or both, will change the specificity. The use of a mutant enzyme may influence the

stability of the perfect matched and mismatched enzyme-DNA complexes to a different extent, so that discrete K_M effects are exerted on these ligation reactions. In a competitive reaction, such as ligation of perfectly matched and mismatched substrates, the ratio of the specificity constant may be altered as a consequence of K_M, and possible k_{cat} changes for each substrate. All mutant enzymes which satisfy the equation below (shown for K294R) will give increased discrimination of cancer-associated mutations in the presence of an excess of normal DNA.

$$\frac{[k_{cat}/K_M]_{K294R,match}}{[k_{cat}/K_M]_{K294R,mismatch}} > \frac{[k_{cat}/K_M]_{Wt,match}}{[k_{cat}/K_M]_{Wt,mismatch}}$$

Alternatively, the second aspect of the present invention can be expressed in terms of a fidelity ratio (i.e. the initial rate of ligating a substrate with a perfect match at the 3' end divided by the initial rate of ligating a substrate with a mismatch at the 3' end) as follows:

$$\frac{[k_1]_{K294R,match}}{[k_1]_{K294R,mismatch}} > \frac{[k_1]_{Wt,match}}{[k_1]_{Wt,mismatch}} = \text{Fidelity ratio}$$

In the above equation, $[k_1]_{match}$ represents the initial rate constant for ligating the first and second oligonucleotide probes hybridized to a target nucleotide sequence with a perfect match at the ligation junction between the target nucleotide sequence and the oligonucleotide probe having its 3' end at the ligation junction. $[k_1]_{mismatch}$ represents the initial rate constant for ligating the first and second oligonucleotide probes hybridized to a target with a mismatch at the ligation junction between the target nucleotide sequence and the oligonucleotide probe having its 3' end at the ligation junction. For the mutant thermostable ligase, $[k_1]_{match}$ divided by $[k_1]_{mismatch}$ (= fidelity ratio) is greater than the fidelity ratio for wild-type ligase. All mutant enzymes which satisfy the equation above (shown for K294R) will give increased discrimination of cancer-associated mutations in the presence of an excess of normal DNA. This can also be stated more generally and in terms of a signal to noise ratio as follows:

$$\frac{[\text{LDR product}]_{\text{minority target}} + [\text{LDR product}]_{\text{majority target}}}{[\text{LDR product}]_{\text{majority target}}} = \text{Signal-to-noise ratio}$$

In the above equation, $[\text{LDR product}]_{\text{minority target}}$ represents the amount of ligation product sequences produced when the first and second oligonucleotide probes hybridize to a minority target nucleotide sequence and the oligonucleotide probe having its 3' end at the ligation junction. $[\text{LDR product}]_{\text{majority target}}$ represents the amount of ligation product sequences produced when the same first and second oligonucleotide probes hybridize to the majority target nucleotide sequence with a mismatch at the ligation junction between the majority target nucleotide sequence and the oligonucleotide probe having its 3' end at the ligation junction. The ligase has a signal-to-noise ratio, for the amount of ligation product sequences produced from both the minority and majority target nucleotide sequences divided by the amount of ligation product sequences produced from the same amount of majority target nucleotide sequences alone.

Both mutant and wild-type ligases have associated signal-to-noise ratios for detection of minority mutations, and the second aspect of the present invention can be expressed as the mutant ligase signal-to-noise ratio is greater than the wild-type ligase signal-to-noise ratio.

$$\frac{\dfrac{[\text{LDR product}]_{\text{minority target}} + [\text{LDR product}]_{\text{majority target}}}{[\text{LDR product}]_{\text{majority target}}}\ \text{Mutant ligase}}{\dfrac{[\text{LDR product}]_{\text{minority target}} + [\text{LDR product}]_{\text{majority target}}}{[\text{LDR product}]_{\text{majority target}}}\ \text{Wild-type ligase}} > 1$$

The above equation may be restated more simply as:

$$\frac{\text{Signal-to-noise ratio for Mutant ligase}}{\text{Signal-to-noise ratio for Wild-type ligase}} > 1$$

For the mutant thermostable ligase, the signal-to-noise ratio is greater than the signal-to-noise ratio for wild-type ligase. All mutant enzymes which satisfy the equation above will give increased discrimination of cancer-associated mutations in the presence of an excess of normal DNA.

The third aspect of the present invention also relates to a method for identifying one or more of a plurality of sequences differing by one or more single-base changes, insertions, deletions, or translocations in a plurality of target nucleotide sequences. As noted above, a sample and one or more oligonucleotide probe sets are blended with a ligase to form a ligase detection reaction mixture. The ligase detection reaction mixture is subjected to one or more ligase detection reaction cycles, and the presence of ligation product sequences is then detected. Here, however, with regard to the oligonucleotide probe sets, the oligonucleotide probe which has its 3' end at the junction where ligation occurs has a modification. This modification differentially alters the ligation rate when the first and second oligonucleotide probes hybridize to a minority target nucleotide sequence in the sample with a perfect match at the ligation junction between the minority target nucleotide sequence and the oligonucleotide probe having its 3' end at the ligation junction compared to the ligation rate when the first and second oligonucleotide probes hybridize to the sample's majority target nucleotide sequence with a mismatch at the ligation junction between the majority target nucleotide sequence and the nucleotide probe having its 3' end at the ligation junction. Ligation with the modified oligonucleotide probe has a signal-to-noise ratio, of the ligation product sequence amounts for the minority and majority target nucleotide sequences to the amount of ligation product sequences produced from the same amount of majority target sequence alone, which is greater than the signal-to-noise ratio for ligation using an oligonucleotide probe lacking the modification.

The use of a modified oligonucleotide probe in accordance with the process of the present invention can be explained as follows:

Introduction of the **Q2** or **Q18** analogues at the 3rd position of the discriminating primer improves the signal to noise ratio about 2 to 3-fold, thereby increasing the power of the LDR system to discriminate cancer signal from background. This assay compares the ability of ligase to discriminate the most difficult case; a T:G mismatch from a T:A perfect match. **Q2** or **Q18** analogues located three nucleotides in from the 3'-end of a probe enhance local melting when present in conjunction with a mismatch at the 3'-position, while at the same time preserving helix integrity more than a mismatch when present in conjunction with a

base pair match at the 3'-end. The use of a **Q2** or **Q18** analogue near the 3' end of a probe may influence the stability of the perfect matched and mismatched enzyme-DNA complexes to a different extent, so that discrete K_M effects are exerted on these ligation reactions. In a competitive reaction, such as ligation of perfectly matched and mismatched substrates, the ratio of the specificity constant may be altered as a consequence of K_M, and possible k_{cat} changes for each substrate. All modified probes which satisfy the equation below (shown for Q analogues) will give increased discrimination of cancer-associated mutations in the presence of an excess of normal DNA.

$$\frac{[k_{cat}/K_M]_{SLP3'QTT,match}}{[k_{cat}/K_M]_{SLP3'QTT,mismatch}} > \frac{[k_{cat}/K_M]_{SLP3'TTT,match}}{[k_{cat}/K_M]_{SLP3'TTT,mismatch}}$$

Alternatively, the third aspect of the present invention can be expressed in terms of a fidelity ratio (i.e. the initial rate of ligating a substrate with an analogue located three nucleotides in from the 3' end as well as a perfect match at the 3' end divided by the initial rate of ligating a substrate with an analogue located three nucleotides in from the 3' end as well as a mismatch at the 3' end) as follows:

$$\frac{[k_1]_{SLP3'QTT,match}}{[k_1]_{SLP3'QTT,mismatch}} > \frac{[k_1]_{SLP3'TTT,match}}{[k_1]_{SLP3'TTT,mismatch}} = \text{Fidelity ratio}$$

The above may be restated more generally to include other nucleotide analogue or sugar phosphate backbone modifications as follows:

$$\frac{[k_1]_{Modified\ oligo,match}}{[k_1]_{Modified\ oligo,mismatch}} > \frac{[k_1]_{Unmodified\ oligo,match}}{[k_1]_{Unmodified\ oligo,mismatch}} = \text{Fidelity ratio}$$

In the above equation, $[k_1]_{\text{Modified oligo, match}}$ represents the initial rate constant for ligating the first and second oligonucleotide probes hybridized to a target nucleotide sequence wherein one oligonucleotide probe contains a modification as well as having a perfect match at the ligation junction between the target nucleotide sequence and the oligonucleotide probe having its 3' end at the ligation junction. $[k_1]_{\text{Modified oligo, mismatch}}$ represents the initial rate constant for ligating the first and second oligonucleotide probes hybridized to a target nucleotide sequence wherein one oligonucleotide probe contains a modification as well as having a mismatch at the ligation junction between the target nucleotide sequence and the oligonucleotide probe having its 3' end at the ligation junction. $[k_1]_{\text{Unmodified oligo, match}}$ represents the initial rate constant for ligating the first and second oligonucleotide probes hybridized to a target nucleotide sequence with a perfect match at the ligation junction between the target nucleotide sequence and the oligonucleotide probe having its 3' end at the ligation junction. $[k_1]_{\text{Unmodified oligo, mismatch}}$ represents the initial rate constant for ligating the first and second oligonucleotide probes hybridized to a target with a mismatch at the ligation junction between the target nucleotide sequence and the oligonucleotide probe having its 3' end at the ligation junction. For the modified oligonucleotide probe, $[k_1]_{\text{Modified oligo, match}}$ divided by $[k_1]_{\text{Modified oligo, mismatch}}$ (= fidelity ratio) is greater than the fidelity ratio for the corresponding unmodified oligonucleotide probe. All modified oligonucleotide probes which satisfy the equation above will give increased discrimination of cancer-associated mutations in the presence of an excess of normal DNA.

Another explanation for the above equation is the oligonucleotide probe which has its 3' end at the ligation junction has one or more modification which differentially alters the rate of ligation when the first and second oligonucleotide probes hybridize to a minority target nucleotide sequence with a perfect match at the ligation junction between the minority target nucleotide sequence and the oligonucleotide probe having its 3' end at the ligation junction, compared to the rate of ligation when the first and second oligonucleotide probes hybridize to the majority target nucleotide sequence with a mismatch at the ligation junction between the majority target nucleotide sequence and the oligonucleotide probe having its 3' end at the ligation junction.

This can also be stated more generally in terms of a signal-to-noise ratio defined as follows:

$$\frac{[\text{LDR product}]_{\text{minority target}} + [\text{LDR product}]_{\text{majority target}}}{[\text{LDR product}]_{\text{majority target}}} = \text{Signal-to-noise ratio}$$

In the above equation, $[\text{LDR product}]_{\text{minority target}}$ represents the amount of ligation product sequences produced when the first and second oligonucleotide probes hybridize to a minority target nucleotide sequence with a perfect match at the ligation junction between the minority target nucleotide sequence and the oligonucleotide probe having its 3' end at the ligation junction. $[\text{LDR product}]_{\text{majority target}}$ represents the amount of ligation product sequences produced when the same first and second oligonucleotide probes hybridize to the majority target nucleotide sequence with a mismatch at the ligation junction between the majority target nucleotide sequence and the oligonucleotide probe having its 3' end at the ligation junction. The ligase, using either modified or unmodified oligonucleotide probes, has a signal-to-noise ratio, for the amount of ligation product sequences produced from both the minority and majority target nucleotide sequences divided by the amount of ligation product sequences produced from the same amount of majority target nucleotide sequence alone.

When using thermostable ligase with both modified and unmodified oligonucleotide probes, there are signal-to-noise ratios associated with each probe for detection of minority mutations, and the third aspect of the present invention can be expressed as the signal-to-noise ratio obtained using modified oligonucleotide probes is greater than the signal-to-noise ratio obtained using unmodified oligonucleotide probes.

$$\frac{[\text{LDR product}]_{\text{minority target}} + [\text{LDR product}]_{\text{majority target}}}{[\text{LDR product}]_{\text{majority target}}}\ \text{Modified Oligo} \;>\; 1$$

$$\frac{[\text{LDR product}]_{\text{minority target}} + [\text{LDR product}]_{\text{majority target}}}{[\text{LDR product}]_{\text{majority target}}}\ \text{Unmodified Oligo}$$

The above equation may be restated more simply as:

$$\frac{\text{Signal-to-noise ratio for modified oligonucleotide}}{\text{Signal-to-noise ratio for unmodified oligonucleotide}} > 1$$

Ligation using the modified oligonucleotide probe has a signal-to-noise ratio, of the amount of LDR product produced from both the minority and majority target nucleotide sequences, to the amount of LDR product produced from the same amount of majority target nucleotide sequence alone, which is greater than the signal-to-noise ratio for ligation using an oligonucleotide probe lacking the modification. All modified oligonucleotide probes which satisfy the equation above will give increased discrimination of cancer-associated mutations in the presence of an excess of normal DNA.

Suitable modifications include the use of nucleotide analogues, such as 1-(2′-Deoxy-β-D-ribofuranosyl)imidazole-4-carboxamide, 1-(2′-Deoxy-β-D-ribofuranosyl)-3-nitropyrrole, 4-(2′-Deoxy-β-D-ribofuranosyl)imidazole-2-carboxamide, 2′-Deoxy-5-fluorouridine, 2′-Deoxyinosine, 6-(2′-Deoxy-β-D-ribofuranosyl)-6H,8H-3,4-dihydropyrimido[4,5-c][1,2]oxazine-7-one, 2-Amino-7-(2′-deoxy-β-D-ribofuranosyl)-6-methoxyaminopurine, 1-(2′-Deoxy-β-D-ribofuranosyl)-5-nitroindole, 1-(2′-Deoxy-β-D-ribofuranosyl)pyrazole-4-carboxamide, 1-(2′-Deoxy-β-D-ribofuranosyl)-1,2,4-triazole-3-carboxamide, 3-Amino-1-(2′-deoxy-β-D-ribofuranosyl)-1,2,4-triazole, 5-(2′-Deoxy-β-D-ribofuranosyl)-2-pyrimidinone, 5-(2′-Deoxy-β-D-ribofuranosyl)-2-thiopyrimidine, 5-Amino-1-(2′-deoxy-β-D-ribofuranosyl)imidazole-4-carboxamide, 1-(2′-Deoxy-β-D-ribofuranosyl)-3-nitropyrazole, 1-(2′-Deoxy-β-D-ribofuranosyl)-4-iodopyrazole, 1-(2′-Deoxy-β-D-ribofuranosyl)-4-propynylpyrazole, 1-(2′-Deoxy-β-D-ribofuranosyl)pyrrole-3-carboxamide, 1-(2'-Deoxy-β-D-ribofuranosyl)pyrazone-4-carboxamide, 1-(2'-Deoxy-β-D-ribofuranosyl)-4-nitroimidazole, or 1-(2′-Deoxy-β-D-ribofuranosyl)-4-nitropyrazole. Alternatively, the modified oligonucleotide probe contains thiophosphate, dithiophosphate, 2'-methoxy, or 3'-amino-2',3'-dideoxy-modifications to the sugar phosphate backbone of the oligonucleotide probe. This modification is either at the position which undergoes ligation, the adjacent position, or the third position for that undergoing ligation.

EXAMPLES

Example 1 - Construction of *Thermus thermophilus* DNA Ligase Mutants at Amino Acid Residue 294 Using Site-Specific Mutagenesis

Thermus thermophilus ("*Tth*") DNA ligase mutants were created using a two-step PCR method (Horton et al., "Engineering Hybrid Genes Without the Use of Restriction Enzymes: Gene Splicing By Overlap Extension," Gene, 77:61-68 (1989), which is hereby incorporated by reference). Plasmid pDZ15 (linearized by *Hin*dIII) was used as a template in the first round of PCR reactions. In the upper panel of Figure 10, plasmid pDZ15 which contains the cloned *Tth* DNA ligase gene (Barany, F., et al., "Cloning, Overexpression, and Nucleotide Sequence of a Thermostable DNA Ligase Gene," Gene, 109:1-11 (1991), which is hereby incorporated by reference), is shown schematically as if opened at a *Pst*I site with genes drawn approximately to scale. In pDZ15, the *Tth* ligase gene and direction is represented by the strongly hatched arrow; the vector Ap^R (*bla*) gene represented by the stippled (gray) arrow; the truncated end of a nonfunctional *Taq* I endonuclease gene represented by the lightly hatched arrow, and the pBR origin of replication represented by the open bar. The *phoA*, and T7 promoters are indicated by right angle arrows and point in the direction of transcription. Restriction sites are: Av, *Avr*II; Bm, *Bam*HI; Bg, *Bgl*II; (Bg/Bm recombined site is not cleavable by either *Bam*HI or *Bgl*II;) Hd, *Hin*dIII; RI, *Eco*RI; Ps, *Pst*I, Pv, *Pvu*II. Polylinker regions from pTZ18R are indicated by the triangular "rake", with only the outside restriction sites listed. *Escherichia coli* host strains used in the constructions described below were obtained from the following sources: N3098 (*lig*ts7; (Wilson, G.G., et al., "Molecular Cloning of the DNA Ligase Gene From the Bacteriophage T4.I. Characterization of the Recombinants," J. Mol. Biol., 132:471-491 (1979), which is hereby incorporated by reference), from N. Murray; JH132 (*mrr*$^-$, Tn10; (Heitman, J., et al., "Site-Specific Methylases Induce the SOS DNA Repair Response in Escherichia coli," J. Bacteriol, 169:3243-3250 (1987), which is hereby incorporated by reference), from J. Heitman; MM294 (*endA*$^-$, *hsdR*$^-$, *hsdM*$^+$, *thi-1, supE44;* (Miller, J.H., "Experiments in Molecular Genetics," Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, pp. 201-205 (1972), which is hereby incorporated by reference), from our collection. Strains AK53 (*mrrB*$^-$, MM294) and AK76 (*mrrB*$^-$, N3098) were constructed by transducing the *mrrB*$^-$ phenotype from JH132 as described (Miller, J.H., "Experiments in Molecular Genetics," Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, pp. 201-205 (1972), which is hereby incorporated by reference).

Presence of the *mrrB*$^-$ phenotype was confirmed by tolerance of these strains to the *Taq* MTase-encoding gene present on plasmid pFBT71 (Barany, F., "A Genetic System for Isolation and Characterization of TaqI Restriction Endonuclease Mutants," Gene, 56:13-27 (1987) and Barany, F., et al., "Cloning and Sequencing of the TthHB8I DNA Restriction and Modification Enzymes, and Comparison With the Isoschizomeric TaqI Enzymes," Gene, 112:3-12 (1992), which is hereby incorporated by reference).

In the lower panel of Figure 10, the *Tth* DNA ligase gene and its direction of transcription is represented by a hatched arrow. The cleavage sites of some restriction endonucleases in the *Tth* DNA ligase gene are indicated by short solid bars. Approximate positions of oligonucleotide primers used for PCR reactions are depicted by arrows and primer names above the *Tth* ligase gene. The site of mutation, amino acid residue K294 is also indicated.

Site-directed mutagenesis at residue K294 was carried out as follows: (1) With *Hind*III-linearized pDZ15 as the template, two independent PCR reactions were performed. In one reaction tube, 400 ng of primers JL505 and JL507R were added to 200 ng of *Hind*III digested pDZ15 containing 50 μmoles of dATP, dCTP, dGTP, and dTTP each, and 2.5 units *Amplitaq* ™ in 100 μl PCR buffer and cycled as described (Saiki, R.K., et al., "Primer-Directed Enzymatic Amplification of DNA With a Thermostable DNA Polymerase," Science, 239:487-491 (1988), which is hereby incorporated by reference); Perkin-Elmer Cetus Instruments, Emoryville CA.) A second reaction tube contained 400 ng of primers JL506 and JL508R, 200 ng of *Hind*III digested pDZ15, and 2.5 units *Amplitaq* ™ in the same reaction buffer, and incubated as above. (2) A third PCR reaction was then carried out in 100 μl PCR buffer containing 1 μl of the products from the two initial reactions, 400 ng primers JL505 and JL508 and 2.5 units *Amplitaq* ™ and incubated as above. After removal of *Amplitaq* ™ the larger product PCR fragment was digested with *Avr*II and *Bam*HI, and electrophoresed in low melting agarose. The 436bp *Avr*II-*Bam*HI fragment was excised from the gel, and purified as described previously (Barany, F., "Overproduction, Purification, and Crystallization of TaqI Restriction Endonuclease," Gene, 63:167-177 (1988), which is hereby incorporated by reference). In addition, plasmid pDZ15 was also digested with *Avr*II and *Bam*HI, and electrophoresed in low melting agarose. The bigger fragment, which equals pDZ15 minus the 436bp *Avr*II-*Bam*HI fragment, was excised and purified. This big fragment was incubated with the 436bp *Avr*II-*Bam*HI fragment purified from the product of the third PCR reaction in the presence of T4 DNA ligase for 16 hours at 14°C. The ligation mixture was transformed into E. coli strain AK76 (ts lig) as described previously (Hanahan, D.,

"Studies on Transformation of E. coli With Plasmids," J. Mol. Biol., 166:557-580 (1983), which is hereby incorporated by reference). Plasmid carrying cells were replica plated onto Fortified Broth plates supplemented with either high concentration of phosphates (10 mM K_2HPO_4, pH 7.6) or low concentration of phosphates (0.2 mM K_2HPO_4, pH 7.6) grown at 32°C and 42°C to test the complementation ability of mutant *Tth* DNA ligase to the *ts lig* host. Independent clones were picked from FB-high phos plates, and grown in liquid Fortified Broth supplemented with 10 mM K_2HPO_4, pH 7.6. Plasmid minipreps were made using the Magic Miniprep kit from Promega, and the *Avr*II-*Bam*HI region was sequenced to confirm site-specific mutations at K294 using the Prism Dye DeoxyTerminator Cycle sequencing kit and DNA sequencer 373A from Applied Biosystems Division of the Perkin-Elmer Corporation.

Example 2 - Construction of *Tth* DNA Ligase Mutants K294R and K294P

The mutant ligases of Figure 11 were prepared using the oligonucleotides of Figure 12A-B as described *infra*. Four oligonucleotide primers for PCR reactions were designed for creating multiple site-specific mutations at K294 site. Their sequences are as follows: Primer a (JL505): 5'CAG AAC CTC CTC ACC ATC 3'; Primer b (JL507R): 5'CTC GTC CAG (G,C) (T, G, C, A) G CAC CAC CAC CCC GTC 3'; Primer c (JL506): 5' TGG TGG TGC (A, C, G, T) (C, G) C TGG ACG AGC TTG CCC T 3'; and Primer d (JL508R): 5' CTC TAT GTA GCT CTC GTT GTG 3'. Primers b and c are overlapping primers containing degenerate codons at mutation site. These primers were synthesized using reagents and a 394 automated DNA synthesizer from Applied Biosystems Division of Perkin-Elmer Corporation, Foster City, CA. After synthesis, primers were deprotected in 30% ammonium hydroxide at 55°C for 12 hours, dried in speedvac, resuspended in 100 µl of ddH_2O (i.e. double distilled water), and purified by ethanol precipitation. The pellet was resuspended in 200 µl of ddH_2O, and their concentrations were determined by spectrophotometry at OD_{260}. Primers were then aliquoted and stored in a -20°C freezer before use.

Site-specific mutagenesis in *Tth* DNA ligase gene (*Tth* lig) was carried out using a two-stage PCR-based overlap extension strategy as described previously (Ho et al., Gene 77:51-59 (1989), which is hereby incorporated by reference). Plasmid pDZ15, the expression plasmid (Barany, et al., Gene 109:1-11 (1991), and Horton, et al., "Engineering Hybrid Genes Without the Use of Restriction Enzymes: Gene Splicing by Overlap Extension," Gene 77:61-68 (1989), which are hereby

incorporated by reference) of *Tth* DNA ligase was linearlized with a restriction endonuclease HindIII, and used as the template for the first round of PCR reactions. Two separate first round PCR reactions were carried out using primers *a* (JL505) and *b* (JL507R) or primers *c* (JL506) and *d* (JL508R), respectively. One microliter of the product from each first round PCR reaction was used as the template for the second round PCR reaction with primers *a* (JL505) and *d* (JL508R). The resulting product was digested with restriction endonucleases *Avr*II and *Bam*HI, and separated on SeePlaque low-melting agarose gel. The DNA fragments assumed to contain the mutation site were cut from the low-melting agarose gel, purified by phenol extraction before ligated to the bigger fragment created from pDZ15 by digestion with AvrII and BamHI. The ligation reaction was carried out at 14°C for 16 hours. The resulting ligation mixture was used to transform AK76, a bacterial strain (*lig ts7* strain, Barany, et al., Gene 109:1-11 (1991), which is hereby incorporated by reference) which contains a temperature-lethal mutation in ligase gene on bacterial chromosome. Positive transformants were selected by growing transformants on TY plates containing 50 μg/ml Ampicillin. Plasmid DNA minipreps were made from the transformants using the Magic Minipreps columns from Promega, and used for sequencing. Regions which was amplified in PCR reactions were sequenced to confirm the mutations using the Prism Dye DeoxyTerminator Cycle sequencing kit and DNA sequencer 373A from Applied Biosystems Division of Perkin-Elmer Corporation, Foster City, CA.

Example 3 - Expression of Mutant *Tth* DNA Ligases in *E. coli*

Plasmids containing mutant *Tth* DNA ligase gene under control of a *pho*A promoter were introduced into *E. coli* strain, AK53 via transformation. Mutant *Tth* DNA ligase proteins were overexpressed at 30°C for 15 hours in 6 ml MOPS medium (Neidhardt, et al., J. Bacteriol., 119:736-47 (1974), which is hereby incorporated by reference) containing 0.2 mM K_2HPO_4 and 75 μg/ml ampicillin (F. Barany, et al., Gene 109:1-11 (1991), which is hereby incorporated by reference). Cells were harvested by centrifugation, resuspended in 400 μl TE, (10 mM Tris, pH 8.5, 1 mM EDTA) sonicated for 3 x 10 seconds with a microprobe on a Sonifier 350 cell disruptor from VWR, and centrifuged for 10 min. at 4°C. The supernatant was adjusted to 20 mM Tris-HCl, pH 8.5, 50 mM KCl, 10 mM $MgCl_2$, and 0.5 mM EDTA, 1 mM DTT, and 2 mM 2-mercaptoethanol. After incubation at 64°C for 25 min, the cloudy suspension was clarified by centrifugation at 4°C for 15 min. Over 70% of the total protein in the resulting clear supernatant is *Tth* DNA ligase, as

determined by staining of a 7.5% polyacrylamide gel containing 0.1% SDS with Coomassie Brilliant Blue. Approximately 200 µg of *Tth* DNA ligase was isolated from a 6 ml culture. Every mutant *Tth* DNA ligase was overexpressed in AK53, and remained soluble after heat treatment at 65°C.

Example 4 - Complementation Assay

Plasmids containing mutated *Tth* DNA ligase genes were introduced into a temperature sensitive ligase defective *E.coli* strain, AK76 (*lig* ts7 strain) (F. Barany, et al., Gene 109:1-11 (1991), which is hereby incorporated by reference) via transformation. Individual transformants were replica plated onto high phosphate plates (0.6% NaCl, 2.5% Bacto-Tryptone, 0.75% yeast extract, 0.1% dextrose, and 10 mM K_2HPO_4, pH 7.6, and 50 µg/ml ampicillin), and low phosphate plates (0.2 mM K_2HPO_4) (Neidhardt, et al., J. Bacteriol., 119:736-47 (1974), which is hereby incorporated by reference); and incubated overnight at permissive (32°C) and non-permissive (42°C) temperatures, respectively. The *Tth* DNA ligase gene is induced only at low phosphate concentration. An active enzyme encoded by the plasmid enables the temperature sensitive host to grow at 42°C on low phosphate plates.

Example 5 - Adenylation Assays

Adenylation was assayed by incubating approximately 8 µg (100 fmoles) wild-type and mutant *Tth* DNA ligase, prepared as described above, in 100 µl of reaction buffer containing 20 mM Tris-HCl, pH 7.6, 50 mM KCl, 10 mM $MgCl_2$, 0.5 mM EDTA, 1 mM NAD^+, 1 mM DTT, 2 mM 2-mercaptoethanol at 65°C for 25 min. Under these conditions, virtually all the wild-type ligase was found in the adenylated form. The reaction was stopped by adding an equal volume of 2 x sample buffer containing 120 mM Tris-HCl, pH 7.6, 2% SDS, 20% glycerol, 0.02% bromophenol blue, and 300 mM 2-mercaptoethanol. Samples were boiled for 5 min, and were analyzed by loading 50 µl on to a 7.5% polyacrylamide-0.1%-SDS gel. The adenylated enzyme can be distinguished by its higher apparent molecular weight (81Kd), compared to the deadenylated form (78Kd).

To rule out the possibility that some mutants may not change mobility after adenylation, experiments were also performed with radioactive [^{32}P] NAD^+. In this case, the reaction was carried out in 25 µl reaction mixture under the same conditions described above except that 3.3 µCi [^{32}P] NAD^+ (800 Ci/mmol, NEN-Du Pont Company, Chadds Ford, PA) was used. After 15 min incubation at 65°C, 1.5

pmoles of non-radioactive NAD^+ was added to the reaction mixture, and incubated for 5 more min to drive the adenylation reaction to near completion. Reactions were stopped by adding 25 µl of 2 x sample buffer. The resulting mixture was boiled at 100°C for 5 min prior to analysis on a 7.5% polyacrylamide - 0.1% SDS gel. The gel was autoradiographed at room temperature for 3 hours against a Kodak XAR-5 film. In order to verify protein samples in each lane, the gel was then stained with Coomassie Brilliant Blue.

Example 6 - Deadenylation Assays

To assay for the deadenylation activity (transfer of the adenyl group from enzyme to DNA substrate), the same conditions were used as for the adenylation experiment, except that 1 mM NAD^+ was replaced by 5 µg of nicked salmon sperm DNA; prepared by incubating salmon sperm DNA (Sigma, St. Louis) with pancreatic DNase I (Barany, et al., Gene 109:1-11 (1991), which is hereby incorporated by reference). The deadenylated enzyme was recognized as a fast migrating band (78 Kd) when separated by electrophoresis on a 7.5% polyacrylamide-0.1%-SDS gel.

Example 7 - 5'-Labeling of Oligonucleotide Probe with [γ-^{32}P] ATP

Oligonucleotide probe JL514 was 5' labeled in a 10 µl reaction containing 50 mM Tris-HCl, pH 7.6, 10 mM $MgCl_2$, 1 mM EDTA, 10 mM DTT, 45 pmole of [g-^{32}P] ATP (6000 Ci/mmol, NEN-Du Pont Company), 15 pmole of gel-purified oligonucleotides, and 10 units of T4 polynucleotide kinase (New England Biolabs, Beverly, MA) after incubation at 37°C for 45 min. 1 µl of 10 mM ATP was added to the reaction mixture, and the incubation was continued for two more minutes. The reaction was quenched by adding 17 µl of 60 mM EDTA. The kinase was heat-inactivated by incubation at 64°C for 20 min. Phosphorylated oligonucleotides were separated on a Sephadex G-25 column equilibrated in TE buffer. Fractions containing phosphorylated oligonucleotides were combined, and stored at -20°C in aliquots before use. The specific radioactivity of phosphorylated oligonucleotide JL514 was 9×10^6 cpm/pmol.

Example 8 - Assay for Nick-Closure Activity

The nicked DNA duplex substrate was formed by annealing two short oligonucleotide probes (JL538 and JL514) to a longer complementary

oligonucleotide target (JL 539). Their nucleotide sequences are: JL538: 5' AAC CAC AGG CTG CTG CGG ATG CCG GTC GGA G 3'; JL514: 5' AGA GCC GCC ACC CTC AGA ACC GCC ACC CTC 3'; JL539: 5' GAG GGT GGC GGT TCT- GAG GGT GGC GGC TCT CTC CGA CCG GCA TCC GCA GCA GCC TGT GGT T 3'. The reaction was carried out in 40 µl of buffer containing 20 mM Tris-HCl, pH 7.6, 10 mM $MgCl_2$, 100 mM KCl, 10 mM DTT, 1 mM NAD^+, and 60 fmol of nicked DNA duplex substrates. DNA probes and target were denatured by incubating the reaction mixture at 94°C for 2 min, and re-annealed at 65°C for 2 min. Ligation reactions were initiated by addition of *Tth* DNA ligase and carried out at 65°C for 30 min. Reactions were terminated by adding 40 µl of stop solution (83% formamide, 8.3 mM EDTA, and 0.17% blue dextran). Samples were denatured at 93°C for 2 min, chilled rapidly on ice prior to loading 20 µl on an 8 M urea-10% polyacrylamide gel. After electrophoresis, the gel was exposed to a phosphorimager screen for 20 min. Radioactively labeled ligation products were analyzed on a Molecular Dynamics Phosphorimager (Sunnyvale, CA) and quantified using Image-Quant software.

The amino acid sequence of the *Tth* DNA ligase gene contains two short sequences, K^{118}VDG and DGVVVK^{294}, which resemble the active site sequence (KYDGQR) of human DNA ligase I. Since human DNA ligase I requires ATP as a cofactor while the *Tth* DNA ligase uses NAD^+ instead, it is possible that their active sites for enzyme-adenylate formation may differ. Although, the sequence K^{118}VDG resembles the active site sequence (KY/LDGXR) of human DNA ligases I, III and IV more than the sequence DGVVVK^{294} does, both sequences were tested for adenylation.

Site-specific mutants were constructed at three amino acid sites, K118, D120, and K294 (Fig. 11). At least four different single amino acid substitutions were made at each site to explore a range of side chain changes (Fig. 11). Mutant *Tth* DNA ligases were overexpressed in AK53 cells, partially purified, and analyzed on a 7.5% polyacrylamide-0.1%-SDS gel (Barany, et al., "Cloning, Overexpression, and Nucleotide Sequence of a Thermostable DNA Ligase Gene," Gene 109:1-11 (1991), which is hereby incorporated by reference). *E. coli* cells transformed with the plasmid vector lacking the *Tth* DNA ligase gene showed no protein in the molecular weight range of *Tth* DNA ligase (76-81Kd), although some heat-resistant impurities from the host bacteria are visible. Wild-type *Tth* DNA ligase and mutant ligases with amino acid substitution at K294 migrate as doublet bands during electrophoresis. The upper band, with an apparent molecular weight of 81Kd is the adenylated form while the lower one at 78Kd is the deadenylated form. Id. Nine out of ten mutant ligases of K118 and D120 migrated as a single band. The exception was D120E, the majority of

which was expressed as the adenylated form while a very small amount of the deadenylated form was seen.

Partially purified mutant *Tth* DNA ligases were assayed for adenylation in the presence of NAD^+, and deadenylation in the presence of nicked salmon sperm DNA. Both wild-type enzyme and all K294 mutants tested became adenylated in the presence of NAD^+ and deadenylated upon incubation with nicked salmon sperm DNA (See Table 1).

Table 1. Effects of amino acid substitutions at K118, D120, and K294 of *Tth* DNA ligase on enzyme activities[a]

Plasmid	Mutant	Adenylation	Deadenylation	Complementation of ts7 lig (in vivo)	Nick-closure activity (%) (in vitro)
pDZ15	Wildtype	+	+	+	100
pJLBE3	K118R	–	NA	–	ND
pJLBE12	K118H	–	NA	–	ND
pJLBE5	K118L	–	NA	–	ND
pJLBE9	K118P1 **(CCC)**	–	NA	–	ND
pJLBE11	K118P2 **(CCG)**	–	NA	–	ND
pJLBg3	D120E	+	±	–	6.2
pJLBF9	D120N	+	±	–	8.5
pJLBF7	D120Y	±	–	–	0.11
pJLBc4	D120G	±	–	–	0.07
pJLBd6	D120A	±	–	–	0.48
pJLBf6	D120V	–	NA	–	ND
pJLBH7	K294R	+	+	+	100
pJLBH2	K294Q	+	+	+	77
pJLBH6	K294L1 **(CTG)**	+	+	+	90
pJLBH10	K294L2 **(CTC)**	+	+	+	87
pJLBH8	K294P	+	+	+	26
pJLBH9	K294P*	+	+	–	ND

[a] Abbreviations: (+), similar activity to wildtype enzyme; (–), no activity; (±), intermediate activity; Deadenylation refers to transfer of the adenyl group from enzyme to nicked DNA substrate; NA, "Deadenylation" could not be determined since these mutants do not adenylate; ND, not detectable; (*), Mutant encoded in plasmid pJLBH9 also contains the G339E second site mutation. For the nick-closure activity experiment, 100% activity represents the formation of 58 fmol product in 30 min. at 65°C using 6 fmol of partially purified wild type *Tth* DNA ligase under conditions described herein. The amount of partially purified mutant enzymes used in this experiment were: 60 fmol for mutants at K118 and D120 sites, 6 fmol for mutants at K294 site, and 60 fmol for mutant K294P* (K294P/G339E). The relative nick-closure activity for mutant *Tth* DNA ligases shown in this Table was normalized based on the amount of enzyme used.

All mutants at K118 site were unchanged by treatments with either NAD^+ or nicked salmon sperm DNA, indicating a possible defect in the adenylation or deadenylation reaction. Likewise, mutant D120V was also defective in adenylation or deadenylation (Table 1). Mutants D120A, D120G, D120Y, and D120N underwent adenylation (see below), although the mobility shifts were difficult to distinguish for some of the

mutants. Most of the mutant D120E protein remained in the adenylated form when isolated from *E. coli* cells, indicating a possible defect during the deadenylation step. When 5 μg of mutant D120E and wild-type enzyme were incubated with 2.5 μg of nicked salmon sperm DNA, wild-type enzyme became completely deadenylated, while most of D120E remained as the adenylated form (data not shown). However, mutant D120E was substantially deadenylated when the amount of nicked salmon sperm DNA was raised above 25 μg. It is unlikely that this change is due to the reversal of the adenylation step since no β-nicotinamide mononucleotide (NMN) was added in the reaction mixture. Thus, mutant D120E either has a reduced affinity to the DNA substrate, or has a reduced rate of transferring the AMP moiety to the 5' phosphate of the DNA substrate.

To rule out the possibility that some mutants may become adenylated without altering mobility on SDS gels, adenylation was also carried out using [^{32}P] NAD^+. None of the K118 mutants were adenylated when assayed with radioactive substrate, while the wild-type enzyme yielded a strong radioactive adenylated-enzyme band. This indicates that K118 is essential for enzyme-adenylate formation. All mutants at D120 site except D120V had incoporated a comparable amount of radioactive AMP relative to that by wild type enzyme. However, when the gel was stained with Coomassie Blue after autoradiography, mutant proteins D120Y and D120G showed only partial adenylation, indicating the higher sensitivity of the radioactive assay. Mutant D120A formed the enzyme-adenylate without changing its electrophoretic mobility, while mutants K118P1 and K118P2 changed mobility not as a result of adenylation, but due to a conformational change caused by the proline for lysine substitution. Thus, a conformational change (as evidenced by altered mobility on an SDS gel) is usually observed upon successful adenylation, but may also be achieved by some of the same mutations which abolish formation of the enzyme-adenylate complex.

The effects of amino acid substitution at K118, D120, and K294 on the overall activity of *Tth* DNA ligase were tested by a complementation assay and an *in vitro* ligation assay (Table 1). In this *in vitro* assay, varying concentrations of wild-type and mutant *Tth* DNA ligases were incubated with a nicked-DNA duplex substrate (composed of two probes hybridized to a synthetic target), and ligation product separated on a denaturing gel. Wild-type *Tth* DNA ligase complemented the *E. coli* lig ts7 host while none of the mutants at K118 and D120 did. All K118 mutants were also defective for *in vitro* ligation activity, but unexpectedly, several D120 mutants retained some *in vitro* activity (Table 1). Mutants D120E and D120N had 6.2% and 8.5% activity respectively, while D120Y, D120G, and D120A all had

less than 0.5% activity. No *in vitro* activity was detected for D120V. All K294 mutants with the exception of one double mutant K294P/G339E, supported the *E. coli* lig ts7 host growth at 42°C, as well as retaining significant *in vitro* enzymatic activity. The aberrant clone did not complement the *E. coli* lig ts7 host, but showed normal activity for adenylation and deadenylation. Sequencing this clone revealed a second mutation of G339E in addition to K294P. The possible involvement of G339 in the formation of phosphodiester bonds was studied further and is discussed below.

The results on single amino acid substitutions of K118, D120, and K294 indicate that K118 is critical for enzyme-adenylate formation, D120 facilitates deadenylation, and K^{118}VDG is thus inferred to be the site of *Tth* DNA ligase-adenylate formation. This supports the prediction from sequence alignment that KXDG may also be the active site of NAD^+-dependent DNA ligases. Similar results using site-directed mutagenesis were reported for ATP-dependent human DNA ligase I (Kodama, K., et al., Nucleic Acids Res., 19:6093-99 (1991), which is hereby incorporated by reference) and vaccinia DNA ligase (Cong, P., et al., J. Biol. Chem., 268:7256-60 (1993) and Shuman, S., et al., Virology, 211:73-83 (1995), which are hereby incorporated by reference). Substitution of the active site Lys (K568) by His or Arg in human DNA ligase I, and of K231 by Asn or Arg in vaccinia DNA ligase caused a loss of the adenylation activity. Mutations at the conserved Asp (D570) to Asn, Glu, and Gln in human DNA ligase I reduced enzyme-adenylate formation and caused loss of *in vivo* complementation (Kodama, K., et al., Nucleic Acids Res., 19:6093-99 (1991), which is hereby incorporated by reference). A KEDG motif was identified as the active site of T4 RNA ligase, based on mass spectrometry of an adenylated peptide (Thogersen, H.C., et al., Eur. J. Biochem., 147:325-29 (1985), which is hereby incorporated by reference), and site-directed mutagenesis studies (Heaphy, S., et al., Biochemistry, 26:1688-96 (1987), which is hereby incorporated by reference). In this enzyme, substitution of the conserved Asp (D101) by Asn, Ser, or Glu was well tolerated for enzyme-adenylate formation, while deadenylation and phosphodiester bond formation steps were prevented completely by each mutation. It was suggested that this Asp residue interacts with the substrate 5'-phosphate terminus rather than the substrate 3'-OH terminus or the adenylate group. All of our D120 substitutions also caused loss of complementation activity (assayed *in vivo* at 42°C), yet D120E and D120N still retained some *in vitro* ligation activity (assayed at 65°C, see Table 1). Therefore, while D120 clearly facilitates deadenylation in *Tth* DNA ligase, it is not strictly essential for ligation. This finding corroborates a similar conclusion by Shuman and Schwer based on studies of capping enzymes and ATP

dependent ligases (Shuman, S., et al., Molec. Microbiol., 17:405-10 (1995), which is hereby incorporated by reference).

A KTDG motif was deduced to be the active site of the mRNA capping enzymes of the Vaccinia virus (Cong, P., et al., J. Biol. Chem., 268:7256-60 (1993), which is hereby incorporated by reference), *S. cerevisiae* (Fresco, L.D., et al., Proc. Natl. Acad. Sci. USA, 91:6624-28 (1994) and Schwer, B., et al., Proc. Natl. Acad. Sci. USA, 91:4328-32 (1994), which are hereby incorporated by reference), and *S. pombe* (Shuman, S., et al., Proc. Natl. Acad. Sci. USA, 91:12046-50 (1994), which is hereby incorporated by reference), for enzyme-guanylate formation. In Yeast tRNA ligase, the amino acid sequence KANG was identified by sequencing the adenylated peptide (Xu, Q., et al., Biochemistry, 29:6132-38 (1990), which is hereby incorporated by reference). A comparison of 5 capping enzymes and 14 ATP dependent DNA and RNA ligases suggests a superfamily of five evolutionarily conserved motifs which plays a role in nucleotidyl binding and transfer to an RNA or DNA substrate (Shuman, S., et al., Molec. Microbiol., 17:405-10 (1995); Shuman, S., et al., Proc. Natl. Acad. Sci. USA, 91:12046-50 (1994); and Cong, P., et al., Molec. Cell. Biol., 15:6222-31 (1995), which are hereby incorporated by reference). These earlier studies, plus the present work on an NAD^+ requiring ligase, allow us to consider KXD/NG as a general active site motif for creating a charged enzyme-nucleotide complex, which provides the energy to form a covalent phosphodiester bond in nucleic acid substrates.

The observation that the double mutant (K294P/G339E) lost ligase activity suggests that G339 may be important for the third step of the ligation reaction; i.e. formation of the phosphodiester bond. To confirm that this effect is caused by one mutation at G339 site, and not by an additive effect of two mutations, single amino acid substitutions were made at G339 by site-directed mutagenesis. Site-specific mutations were also made at R337, a conserved positively charged amino acid near G339, and at C412, C415, C428, and C433 (Fig. 11). There are only four Cys residues in *Tth* DNA ligase, all conserved among the five NAD^+-dependent bacterial ligases that are sequenced. These four Cys residues may form a zinc-binding site and be involved in the interaction between bacterial DNA ligase and DNA substrates (Thorbjarnardottir, S.H., et al., Gene, 161:1-6 (1995), which is hereby incorporated by reference).

All *Tth* DNA ligase mutants constructed at these six sites were able to form the enzyme-adenylate complex in the presence of NAD^+, and were deadenylated in the presence of nicked salmon sperm DNA (Table 2). The effects on the overall

ligase activity varied with each mutation, but was generally consistent, when comparing *in vivo* and *in vitro* activities (Table 2).

Table 2. Effects of single amino acid substitutions at R337, G339, C412, C415, C428, C433 of *Tth* DNA ligase on enzyme activity[a]

Plasmid	Mutant	Adenylation	Deadenylation	Complementation of ts7 lig (in vivo)	Nick-closure activity (%) (in vitro)
pDZ15	Wildtype	+	+	+	100
pJLBK8	R337K	+	+	+	3.1
pJLBK4	R337Q	+	+	–	0.70
pJLBK5	R337E	+	+	–	0.04
pJLBJ6	G339A	+	+	–	0.71
pJLBJ5	G339D	+	+	–	0.34
pJLBA6	G339E	+	+	–	0.22
pJLBB4	C412A	+	+	+	41
pJLBB9	C412V	+	+	–	0.16
pJLBB11	C412T	+	+	–	0.11
pJLBB12	C412M	+	+	–	0.17
pJLBC18	C415A	+	+	+	79
pJLBC3	C415V	+	+	+	100
pJLBC9	C415T	+	+	+	3.0
pJLBC19	C415M	+	+	+	0.47
pJLBD10	C428A	+	+	+	57
pJLBD6	C428T	+	+	–	0.43
pJLBI11	C433A	+	–	–	0.15
pJLBI6	C433V	+	–	–	0.02
pJLBI4	C433T	+	–		0.07
pJLBI1	C433M	+	–		0.06

[a] Abbreviations: (+), similar activity to wildtype enzyme; (–), no activity; ND, not detectable. For the nick-closure activity experiment, 100% activity represents the formation of 58 fmol product in 30 min. at 65 °C using 6 fmol of partially purified

wild type *Tth* DNA ligase or the formation of 16.2 fmol product in 30 min at 65°C using 0.6 fmol. The amount of the other enzymes used was either 0.6 fmol, 6 fmol, or 60 fmol. The relative nick-closure activity for mutant *Tth* DNA ligases shown in this Table was normalized based on the amount of enzyme used.

Mutants R337Q and R337E lost activity and were unable to complement *E. coli* lig ts7 host, while mutant R337K retained partial *in vitro* activity and complementation activity. Substitution of G339 by Ala, Asp, or Glu all rendered the enzyme inactive, both *in vivo* and *in vitro*. At C412, substitution by Ala had no effect in both *in vivo* and *in vitro* experiments, while substitution by Val, Thr, and Met caused the loss of the overall enzyme activity. Similar results were also observed at C428. Three of four mutations at C415 site (C415A, C415V, and C415T), all retained complementation and enzymatic activity. It is not clear why C415M retained complementation activity (assayed at 42°C) with such poor enzymatic activity (assayed at 65°C), although the mutation may have rendered the enzyme thermolabile. In contrast, all mutations at C433 site (C433A, C433V, C433T, or C433M) caused loss of both complementation and enzymatic activity.

While the adenylation active site in many DNA ligases has been well defined, the possible active site for formation of the phosphodiester bond remains poorly understood. Residues G339 and C433 may be involved in this third step of the ligation reaction, because conservative mutations at both sites abolished the overall enzyme activity without affecting the first two steps, adenylation and deadenylation. Residue G339 may allow a local structure critical for enzymatic activity, or interact with the DNA substrate via the peptide backbone, in a manner which would be incompatible with an Ala, Glu, or Asp substitution. Glycine residues play essential roles in mRNA capping enzymes, although these mutations interfered primarily with enzyme-guanylate formation (Shuman, S., et al., Proc. Natl. Acad. Sci. USA, 91:12046-50 (1994) and Cong, P., et al., Molec. Cell. Biol., 15:6222-31 (1995), which are hereby incorporated by reference). Residues R337, C412, and C428 may play an indirect, but not essential, role in the third step of ligation reaction since only nonconservative mutations caused loss of activity for this step. For all mutants at these five residues, there were no dramatic global conformational changes induced by these mutations as indicated by their ability to form enzyme-adenylate complex in the presence of NAD^+, and to deadenylate in the presence of nicked DNA substrate. The active site for the third step of ligation may well be separated from that of the first two

steps. These active sites function independent of one another and damage at one site may not affect the activity of the other.

In summary, site-directed mutagenesis studies were carried out to identify active sites for *Tth* DNA ligase. The adenylation active site and facilitated deadenylation site of *Tth* DNA ligase was identified as the Lys and Asp residues of motif $K^{118}VDG$. These results are consistent with other mutagenesis studies on the active sites of DNA ligases (Kodama, K., et al., Nucleic Acids Res., 19:6093-99 (1991); Cong, P., et al., J. Biol. Chem., 268:7256-60 (1993); and Shuman, S., et al., Virology, 211:73-83 (1995), which are hereby incorporated by reference), T4 RNA ligase (Heaphy, S., et al., Biochemistry, 26:1688-96 (1987), which is hereby incorporated by reference) and mRNA capping enzymes (Cong, P., et al., J. Biol. Chem., 268:7256-60 (1993); Fresco, L.D., et al., Proc. Natl. Acad. Sci. USA, 91:6624-28 (1994); Schwer, B., et al., Proc. Natl. Acad. Sci. USA, 91:4328-32 (1994); Shuman, S., et al., Proc. Natl. Acad. Sci. USA, 91:12046-50 (1994); and Cong, P., et al., Molec. Cell. Biol., 6222-31 (1995), which are hereby incorporated by reference). Together, they support the idea that KXD/NG is the active site motif for the formation of enzyme-nucleotide complex. Mutations at residues G339 and C433 did not inhibit adenylation and deadenylation steps but abolished the overall activity, indicating that these two amino acid residues may be involved in the third step of ligation, the formation of the phosphodiester bond.

Example 9 - Synthesis of Oligonucleotide Probes

Oligonucleotide probes were synthesized using reagents and a model 394 automated DNA synthesizer from Applied Biosystems Division of Perkin-Elmer Corporation, (Foster City, CA). Fluorescent label was attached to the 5' end of oligonucleotides using 6-FAM (6-carboxy fluorescein) amidites, or attached to a 3'-amino group (C3-CPG column from Glen Research (Sterling, VA)) using NHS-FAM (N-hydroxysuccinimide ester of FAM) from Applied Biosystems Division of Perkin-Elmer Corporation. A universal nucleotide analogue, 1-(2'-deoxy-b-D-ribofuranosyl)-3-nitropyrrole, herein designated as **Q**, was synthesized, transformed to the phosphoramidite, and oligonucleotides synthesized as described (Bergstrom, D.E., et al., JACS, 117:1201-1209 (1995), which is hereby incorporated by reference). All oligonucleotides used in this study were purified by polyacrylamide gel electrophoresis with recovery of DNA from gel slices using C-18 Sep-Pak Cartridges from Waters Division of Millipore (Bedford, MA).

Example 10 - 5'-Phosphorylation of Oligonucleotide Probes

One nmole of gel-purified oligonucleotide was phosphorylated in a 25 µl reaction containing 50 mM Tris-HCl, pH 7.6, 10 mM $MgCl_2$, 1 mM EDTA, 10 mM DTT, 1 mM ATP, and 10 units of T4 polynucleotide kinase (New England Biolabs, Beverly, MA) at 37°C for 45 minutes. The reaction was quenched by adding 0.5 µl of 0.5 M EDTA, and the kinase was heat-inactivated by incubation at 64°C for 20 minutes. The phosphorylated oligonucleotides were stored at -20°C in 5 µl aliquots before use.

Example 11 - Purification of Wild-Type and Mutant *Tth* DNA Ligase

Wild-type *Tth* DNA ligase was purified from an *E.coli* strain containing the *Tth* ligase gene under *pho*A promoter control as described (Barany, F., et al., Gene, 109:1-11 (1991)), which is hereby incorporated by reference) with some modifications. Briefly, cells were grown overnight at 30°C in low phosphate (inducing) medium, harvested, resuspended in lysis buffer (20 mM Tris-HCl pH 8.5, 1 mM EDTA, 10 mM 2-mercaptoethanol, 0.15 mM PMSF) and sonicated (Barany, F., et al., Gene, 109:1-11 (1991), which is hereby incorporated by reference). After removal of cellular debris, the supernatant was adjusted to 20 mM Tris-HCl pH 8.5, 50 mM KCl, 10 mM $MgCl_2$, 0.5 mM EDTA, 1 mM DTT, and 2 mM 2-mercaptoethanol, incubated at 65°C for 30 min., and cleared by centrifugation at 4°C. The supernatant was diluted with an equal volume of 10 mM Tris-HCl, pH 7.6, 0.5 mM EDTA, filtered, and loaded onto a 10 ml Red-Sepharose column (Pharmacia) equilibrated with 20 mM Tris-HCl, pH 8.5, 50 mM KCl, 1 mM EDTA, 20% glycerol. The protein was eluted with a 30 ml linear salt gradient of 50 mM to 1 M KCl (Takahaski, M., et al., Agric. Biol. Chem., 50:1333-1334 (1986), which is hereby incorporated by reference), using an FPLC apparatus from Pharmacia. *Tth* DNA ligase eluted between 0.4 - 0.8 M KCl, and fractions containing pure *Tth* DNA ligase (seen as a doublet of adenylated and deadenylated forms on Coomassie-stained 7.5% polyacrylamide-0.1% SDS polyacrylamide gels), were pooled. The enzyme was precipitated with an equal volume of saturated ammonium sulfate, the pellet dissolved in 1.5 ml dH_2O, and dialyzed at 4°C against 500 ml storage buffer containing 10 mM Tris-HCl, pH 8.5, 1 mM EDTA, 1 mM DTT, 0.2 mg/ml BSA, and 50% glycerol. Protein concentration was determined by the Bradford method (Bradford, M.M., Anal. Biochem., 72:248-254 (1976), which is hereby incorporated by reference).

About 4 mg of *Tth* DNA ligase was obtained from 450 ml culture. Mutant *Tth* DNA ligase was partially purified as described previously.

Example 12 - Fidelity Assays of Nick Closure by *Tth* DNA Ligase

Each reaction was performed in 40 µl of buffer containing 20 mM Tris-HCl, pH 7.6; 10 mM $MgCl_2$; 100 mM KCl; 10 mM DTT; 1 mM NAD^+; and 12.5 nM (500 fmoles) of nicked DNA duplex substrates. DNA probes were denatured (94°C for 2 min), re-annealed (65°C for 2 min), and ligations initiated by the addition of 0.125 nM (5 fmoles) *Tth* DNA ligase and carried out at 65°C. Five µl aliquots were removed at 0 hr, 2 hr, 4 hr, 6 hr, 8 hr, and 23 hr, and mixed with 18 µl of a stop solution (83% formamide, 8.3 mM EDTA, and 0.17% Blue Dextran). To 5 µl of this mixture, 0.5 µl of ROX-1000, a fluorescently labeled in-lane size standard (Applied Biosystems Division of Perkin-Elmer) was added. Samples were denatured at 93°C for 2 min, rapidly chilled on ice prior to loading on an 8 M urea-10% polyacrylamide gel, and electrophoresed at 1400 V (constant voltage) on a model 373A automated DNA Sequencer (Applied Biosystems Division of Perkin-Elmer Corporation). Electrophoresis conditions were modified as suggested by the manufacturer. The gel was polymerized in 1.2 x TBE (54 mM Tris-Borate and 1.2 mM EDTA, pH 8.0) and was pre-run before loading samples in a running buffer of 0.6 x TBE (27 mM Tris-Borate and 0.6 mM EDTA, pH 8.0) for 30 min with an electrode polarity opposite to the normal run with samples. After pre-run and sample-loading, the gel was run in 0.6 x TBE in the normal top to bottom direction for 2.5 hrs. Fluorescently labeled ligation products were analyzed and quantified using Genescan 672 version 1.2 software (Applied Biosystems Division of Perkin-Elmer Corporation), and the results were plotted using DeltaGraph Pro3 software (DeltaPoint, Inc. Monterey, CA).

Example 13 - Measurement of Initial Rates of Perfect Match and Mismatch Ligations by *Tth* DNA Ligase

Conditions for these experiments were the same as that for the fidelity assay except that different amounts of *Tth* DNA ligase and different probes were used, as indicated in Brief Description of the Drawings. Reactions were carried out in 160 µl of reaction buffer containing 12.5 nM (2 pmoles) of nicked DNA duplex substrates at 65°C. DNA probes and target were denatured by incubating the reaction mixture at 94°C for 2 min, and re-annealed at 65°C for 2 min. Ligations were initiated by the addition of the *Tth* DNA ligase. Aliquots (10 µl) were removed at 0, 2, 4, 6, 8, and 10

hr for reactions containing mismatched substrates; and at 0, 10, 20, 30, 40, 50, 60, and 70 sec for reactions containing matched substrates. For assays using matched substrates, 2.5 µl samples were mixed with 2.5 µl loading buffer, 0.5 µl of ROX-1000 before gel electrophoresis. Since the linear detection range of fluorescent samples on the 373A DNA Sequencer is from 0.1 fmol to 10 fmol, products from mismatch ligation with a yield less than 1% were concentrated by ethanol precipitation for accurate quantification. From 10 µl aliquot, 9.5 µl was brought up to 200 µl with TE buffer (10 mM Tris-HCl, pH 8.0, 1 mM EDTA) and ethanol precipitated with 4 µg of Yeast tRNA as carriers. The pellet was resuspended in 5 µl loading buffer and 0.5 µl ROX-1000 before gel electrophoresis. The amount of unreacted fluorescent probe was determined by diluting 0.5 µl of the 10 µl aliquot with 4.5 µl loading buffer plus 0.5 µl ROX-1000. Samples were separated on denaturing polyacrylamide gels, and results analyzed as described above. The initial rates were calculated as the slope of the straight line in the graph with the x-axis as the time, and the y-axis as the yield of products.

Example 14 - Preparation of Oligonucleotide Probes Containing Base Analogues and Mismatch in Third Position on the 3' Side of the Nick

Oligonucleotide probes were designed to test the possibility of improving the fidelity of LDR reactions using the *Tth* DNA ligase by introducing a base analogue or a mismatch at the third position on the 3'-side of the nick. These ten descriminating probes were made in 5 pairs. The two probes in each pair differ by one base at the 3' end ("C" or "T"). The probe with the "C" at the 3' end has two more bases at its 5' end than the "T" probe, allowing one to distinguish the ligation products on a sequencing gel. The base analogue "Q" and other bases at the third position from the 3' end are underlined. A nicked DNA duplex substrate is formed by annealing one of left side probes, (for example, SLP3'TTT), and the common fluorescently labeled probe Com 610-3'F to a template probe (for example, GLg.m3A). This substrate contains a T-G mismatch on the 3' side of the nick. Accurate quantification of mismatch and perfect match ligation products can be achieved by scanning the fluorescently labeled products using Genescan program. A ratio of initial rates of perfect match ligation over mismatch ligation indicates the fidelity of the LDR reaction. An extra mismatch or a base analogue was introduced in the third position on the 3' side of the nick in order to improve the fidelity. In Figure 17A, probes SLP3'Q_2TC, SLP3'Q_2TT, SLP3'TTC, and SLP3'TTT were used on targets GLg.m3A to test the contribution of Q_2:A to increased fidelity compared with a T:A match at

the 3rd position from the 3'-side. The SLP3'TTC and SLP3'TTT probes were used on mixtures of targets GLg.m3A and ALg.m3A to test for detecting a rare target (cancer mutation) in the excess of a common target (normal DNA). In Figure 13B, probes SLP3'ATC and SLP3'ATT were used on targets GLg to test the contribution of A:C to increased fidelity compared with a G:C match at the 3rd position from the 3'-side. In Figure 17C, probes SLP3'Q_2TC, SLP3'Q_2TT, SLP3'Q_{18}TC, SLP3'Q_{18}TT, SLP3'GTC, and SLP3'GTT were used on target GLg.m3T to test the contribution of Q_2:T, Q_{18}:T or G:T to increased fidelity compared with a T:A match at the 3rd position from the 3'-side.

Example 15 - Fluorescent Assay

As shown in Figures 13A-C, one of the four long oligonucleotides GLg, ALg, TLg or CLg (shown in the Figure 13C) was used as the template strand, which vary at the underlined base. Figures 13A-B represent the formation of nicked DNA duplex using one of the template strands, ALg, as an example. Shown in Figure 13A, 4 different nicked DNA substrates are formed by annealing the common fluorescently labeled oligo, com5F, and one of the discriminating oligos (RP5'A, RP5'C, RP5'G, RP5'T) to the template strand, ALg. In Figure 13B, 4 different nicked DNA substrates are formed by annealing the fluorescently labeled oligo, com3F, and one of the discriminating oligos (LP3'A, LP3'C, LP3'G, LP3'T) to the template strand, ALg. A matrix of nicked DNA duplexes is thus formed with all possible combinations of match and mismatch base pairing on the 3' and the 5' side of the nick, when ALg is replaced by one of the following template strands, GLg, TLg, and CLg. Products formed by ligation to the common fluorescently labeled probes can be discriminated by size on denaturing polyacrylamide gel due to the incorporation of different length of "A " tails.

Sequences of these probes (shown in Figure 14) were derived from that of human eukaryotic protein synthesis initiation factor eIF-4E (Rychlik, W., et al., "Amino Acid Sequence of the mRNA Cap-Binding Protein From Human Tissue," Proc. Natl. Acad. Sci. USA, 84:945-949 (1987), which is hereby incorporated by reference). A random DNA sequence from a eukaryotic source was chosen to avoid any false signal arising from possible bacterial DNA contamination in *Tth* DNA ligase preparation. The melting temperature of probes were predicted using the nearest neighbor thermodynamic method (Breslauer, K.J., et al., "Predicting DNA Duplex Stability From the Base Sequence," Proc. Natl. Acad. Sci. USA, 83:3746-3750 (1986), which is hereby incorporated by reference). OLIGO 4.0 program from

National Biosciences Inc., Plymouth, MN was used to rule out possible hairpin structure, repetitive sequences, and false priming. The templates and detecting oligonucleotides for this assay have been designed such that their melting temperature is sufficiently higher than the temperature used in the assay (65°C) to minimize the effect of the melting temperature of probes on ligation efficiency.

All oligonucleotide probes were synthesized using reagents and a 394 automated DNA synthesizer from Applied Biosystem Division of Perkin-Elmer Corporation, Foster City, CA. Synthesis of oligonucleotides with a fluorescent dye, 6-FAM (6-carboxy Fluorescin), attached at the 5' end was done using 6-FAM Amidites from Applied Biosystem Division of Perkin-Elmer Corporation. The oligonucleotide with a 3' FAM was made by using a 3'-Amino-modifier C3-CPG column from Glen Research (Sterling, VA) for the initial DNA synthesis, and the FAM group was then attached through the 3'-amino group using NHS-FAM (N-hydroxyl Succinimide ester of FAM) from the Applied Biosystem Division of Perkin-Elmer Corporation. All oligonucleotides used in this study were purified by polyacrylamide gel electrophoresis and recovery of DNA from gel slices using C-18 Sep-Pak Cartridges from Waters Division of Millipore.

Oligonucleotide probes, RP5'A, RP5'C, RP5'G, RP5'T, and Com3'F were 5'-phosphorylated in a solution containing 50 mM Tris-HCl, pH 7.6, 10 mM $MgCl_2$, 1 mM EDTA, 10 mM DTT, 1 mM ATP, 1 nmole of gel-purified oligonucleotides and 10 units of T4 polynucleotide kinase (New England Biolabs, Beverly, MA) at 37°C for 45 minutes. The reaction was stopped by adding 0.5 µl of 0.5 M EDTA, and the kinase was heat-inactivated by incubation at 64°C for 20 minutes. The phosphorylated oligonucleotides were stored at -20°C in aliquots before use.

The fluorescent fidelity assay of *Tth* DNA ligase was carried out in 40 µl of buffer containing 20 mM Tris-HCl, pH 7.6; 10 mM $MgCl_2$; 100 mM KCl; 10 mM DTT; 1 mM NAD^+; and 500 fmol (12.5 nM) of nicked duplex substrates. DNA probes were denatured by incubating the reaction mixture at 94°C for 2 min, and re-annealed at 65°C for 2 min. Ligations were started by the addition of 5 fmol of the thermostable ligase and carried out at 65°C. 5 µl aliquots were removed at 0 hr, 2 hr, 4 hr, 6 hr, 8 hr, and 23 hr, and mixed with 18 µl of a stop solution (83% Formamide, 8.3 mM EDTA, and 0.17% Blue Dextran). 5 µl of this mixture was denatured at 93°C for 2 min, chilled rapidly on ice prior to loading on an 8 M urea-10% polyacrylamide gel, and electrophoresed on a 373A automated DNA Sequencer from Applied Biosystem Division of Perkin-Elmer Corporation, Foster City, CA. Fluorescently labeled ligation products were analyzed and quantitated using Genescan 672 software

(Applied Biosystem Division of Perkin-Elmer Corporation, Foster City, CA). Genescan 672 software provided analyzed data in the form of a gel image and a table containing the peak height and peak area of each peak in each lane. Typically, two bands were seen in each lane representing each reaction. The lower one was the free fluorescent common oligonucleotide, the upper one was the upper strand of the ligation product. The product yield in percentage was calculated as product over total intial substrates times 100. The amount of product was calculated as the peak area of the appropriate ligation product. The amount of initial substrates were calculated by adding the peak area of the product peak to that of the free fluorescent oligonucleotide peak. Results were plotted using DeltaGraph Pro3 software (DeltaPoint, Inc. Monterey, CA) with time as the X-axis and yield as the Y-axis.

Strategy for testing wild-type *Tth* DNA ligase fidelity

A fluorescent assay using nicked substrates was developed for testing *Tth* DNA ligase fidelity. The nicked duplex substrate was generated by annealing two adjacent oligonucleotide probes (one of them containing a fluorescent dye FAM) to a longer complementary template (bottom) strand (see Figure 13 for sequences). For clarity, the fluorescently labeled probe is defined as the common oligonucleotide while the probe containing the test base at its terminus is called the discriminating oligonucleotide. A set of 14 oligonucleotides (Figure 13) were used to generate all possible combinations of different single-base pair matches and mismatches at the 3'- and 5'-sides of the nick. Both common and discriminating oligonucleotides were designed such that their melting temperature was at least 10°C higher than the assay temperature (65°C). This presumably minimized the effect of differences in oligonucleotide hybridization on ligation efficiency. The ligation time was extended to 23 hours allowing for accurate quantification of mismatch ligation products. Ligation of the two adjacent oligonucleotides formed a longer fluorescent product, which was separated on a denaturing polyacrylamide gel and analyzed.

Fidelity of nick closure by wild-type *Tth* DNA ligase

A time course for nick closure by *Tth* DNA ligase using substrates where the discriminating base is on the 3'-side of the nick is recorded in Table 3 and is shown in Figure 15.

Table 3: Ligation yield generated by *Tth* DNA ligase with different DNA substrates containing different baseparing on the 3' side of the nick.

Base-pairing	0 hour	2 hour	4 hour	6 hour	8 hour	23 hour
A-A	0%	0%	0%	0%	0%	0%
A-T	0%	85.3%	88.9%	90.2%	92%	93.2%
A-G	0%	0%	0%	0%	0%	0%
A-C	0%	0%	0%	0%	0%	1.1%
C-A	0%	0%	0%	1.3%	1.8%	2%
C-T	0%	0%	0%	0%	0%	0%
C-G	0%	82%	88%	90%	96.5%	93.3%
C-C	0%	0%	0%	0%	0%	0%
G-A	0%	0%	0%	0%	0%	0%
G-T	0%	2.5%	3.8%	5.46%	7.5%	11%
G-G	0%	0%	0%	0%	0%	0%
G-C	0%	82.6%	87.4%	87.4%	88.9%	91.3%
T-A	0%	83.8%	86.5%	88.4%	89%	92.6%
T-T	0%	0%	0%	0%	0%	0%
T-G	0%	2.2%	3.8%	4.45%	8.8%	13.4%
T-C	0%	0%	0%	0%	0%	0%

Each panel in Figure 15 shows the yield of product formed with the same discriminating oligonucleotide and common oligonucleotide annealed to four template strands which differ by a single base. All perfectly matched substrates yielded over 80% product within 2 hrs. (Fig. 15). Of all 12 mismatches tested on the 3'-side of the nick, T-G and G-T mismatches were less efficiently discriminated, with yields of about 2% after 2 hrs. accumulating to about 15% after 23 hrs. incubation, (Fig. 15). These results on discriminating different 3' side mismatches with *Tth* DNA ligase are similar to those reported for T4 DNA ligase (Landegren, U., et al., Science, 241:1077-1080 (1988), which is hereby incorporated by reference), although *Tth* DNA ligase does not require high salt, spermidine, or very low enzyme concentrations to suppress mismatch ligations (Wu, D.Y., et al., Gene, 76:245-254 (1989) and

Landegren, U., et al., Science, 241:1077-1080 (1988), which are hereby incorporated by reference).

When the mismatches were located at the 5'-side of the nick, the enzyme still exhibited stringent discrimination against A-G, C-C, G-G, and T-C mismatches. The results of this experiment are recorded in Table 4 and are plotted in Figure 16.